Essays in Empirical Macroeconomics: Identification in Vector Autoregressive Models and Robust Inference in Early Warning Systems

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
doctor rerum politicarum
(Dr. rer. pol.)

am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Martin Bruns

Berlin, 2019

Erstgutachter:	Prof. Dr. Helmut Lütkepohl
	Freie Universität Berlin und DIW Berlin
Zweitgutachter:	Prof. Dr. Dieter Nautz
	Professur für Ökonometrie
	Freie Universität Berlin

Disputation: 1. Juli 2019

To Sarah

Acknowledgments

First, I would like to thank Helmut Lütkepohl for excellent supervision, for providing the guidance and advice I needed as well as the freedom to build my own research agenda. In addition to shaping this dissertation and giving valuable feedback, I am grateful for all the help I received with applications for scholarships, conferences, and research positions. My second supervisor, Dieter Nautz, provided great help and support, especially during the job market, and gave constructive comments during numerous seminars and meetings.

My colleagues and friends at the DIW Graduate Center were a great support during countless discussions over lunch, coffee or outside the office. Niels Aka, Khaled el-Fayoumi, Jakob Miethe and Julia Schmieder were great companions during the whole PhD. In addition, my office mates Pablo Anaya, Daniel Bierbaumer, Caterina Forti-Grazzini, Sandra Pasch and Thore Schlaak gave valuable advice, feedback and support. The GC Team, in particular Georg Weizsäcker, Juliane Metzner and Daniela Centemero, provided support throughout the PhD.

All participants at the Freie Universität doctoral seminar provided helpful comments, and helped improve all projects of this dissertation.

My co-author Michele Piffer helped me appreciate the potential of Bayesian econometrics and taught me how fruitful joint work on common projects can be. From my co-author Tigran Poghosyan I learned how to attract the interest of policy makers and I learned about the work at the International Monetary Fund. I am hoping for many collaborations in the future.

This dissertation would have been impossible without the help of the German Studienstiftung for which I am grateful. They provided financial support as well as the opportunity to participate in inter-disciplinary research seminars, language courses and other events.

Above all, I am grateful to my parents and to Sarah Playfair for support and patience during the last years.

Berlin, Mai 2019 *Martin Bruns*

Erklärung zu Ko-Autorenschaften

Diese Dissertation besteht aus drei (Arbeits-)Papieren, von denen zwei in Zusammenarbeit mit einem Koautor entstanden sind. Der Eigenanteil an Konzeption, Durchführung und Berichtsabfassung der Kapitel lässt sich folgendermaßen zusammenfassen:

- Martin Bruns und Tigran Poghosyan:

  "Leading Indicators of Fiscal Distress: Evidence from Extreme Bounds Analysis"

  Eigenanteil: 50 Prozent

- Martin Bruns und Michele Piffer:

  "Bayesian structural VAR models: a new approach for prior beliefs on impulse responses"

  Eigenanteil: 50 Prozent

- Martin Bruns:

  "Combining Factor Models and External Instruments to Identify Uncertainty Shocks"

  Eigenanteil: 100 Prozent

Liste der Vorpublikationen

Publikationen in referierten Fachzeitschriften

Bruns, Martin, and Tigran Poghosyan. Leading indicators of fiscal distress: evidence from extreme bounds analysis. Applied Economics 50.13 (2018): 1454-1478.
Vorpublikation von Kapitel 1

Working Papers

Bruns, Martin and Michele Piffer. Bayesian Structural VAR Models: A New Approach for Prior Beliefs on Impulse Responses. DIW Discussion Papers 1796, Berlin 2019
Vorpublikation von Kapitel 2

Bruns, Martin, and Michele Piffer. Bayesian Structural VAR models: a new approach for prior beliefs on impulse responses. No. 878. Queen Mary University of London, School of Economics and Finance, 2018.
Vorpublikation von Kapitel 2

Contents

List of Figures

List of Tables

List of Abbreviations

AR	auto-regression
APD	absolute percentage deviation
BMA	Bayesian model averaging
BP-FAVAR	Bayesian proxy factor-augmented vector auto-regression
BP-VAR	Bayesian proxy vector auto-regression
BVAR	Bayesian vector autoregression
CE	cross-entropy
CPI	consumer price index
DSMH	dynamic striated Metropolis-Hastings
EBA	extreme bounds analysis
EWS	early warning system
ESS	effective sample size
FAVAR	factor augmented vector autoregressive
FDI	foreign direct investment
Fed	Federal Reserve System
FRED	federal reserve economic data
FX	foreign exchange
GDP	gross domestic product
H_0	null hypothesis
IMF	International Monetary Fund
IP	industrial production
IS	importance sampling
LIBOR	London interbank offered rate
LR	likelihood ratio
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
NiWU	normal inverse-Wishart uniform
Np(B)	normal p(B)
OLS	ordinary least squares

PC	principal components
SVAR	structural vector auto-regression
S&P	Standard and Poor's
TBill	treasury bill
TME	total misclassification error
US	United States
VAR	vector auto-regression
VXO	Central Bank of England S&P 100 volatility index
ZLB	zero lower bound

Summary

Bayesian empirical macroeconomic models are excellent tools for prediction and structural analysis. The use of a prior distribution facilitates model averaging, allows for structural identification of multiple time series models and makes estimation of high-dimensional models feasible. However, prior distributions need to be chosen carefully in order to accurately reflect the researcher's beliefs before looking at the data. I exemplify how to do so in this thesis by employing model averaging techniques in Bayesian spirit, by developing tools to express priors for structural vector-autoregressive models, and by showing a new approach to identify the impact of variations in uncertainty in a data-intensive environment.

In the first chapter, which is based on joint work with Tigran Poghosyan, we use an Early Warning System (EWS) to recover leading indicators of fiscal distress events. In particular, we use Extreme Bounds Analysis (EBA), a model averaging approach introduced by Leamer (1985) and popularised by Sala-i Martin (1997), to assess the robustness of leading indicators for fiscal distress across different models. We find that both fiscal and non-fiscal leading indicators are robust predictors of fiscal distress events. In a second step we assess the forecasting performance of an EWS based on the most robust leading indicators. We find that it offers a gain in predictive power compared to a baseline model which is based on fiscal leading indicators only. Lastly, we assess the robustness of these results across various model specifications, subsamples and estimation strategies.

In the second chapter, which is based on joint work with Michele Piffer, we propose a new approach to express prior beliefs on the impulse responses of structural vector auto-regressive (SVAR) models. This approach does not restrict the family of prior distributions to a set that is technically convenient. Rather, it combines extensive flexibility in the choice of priors with an efficient importance sampler to explore the posterior distribution. We compare our new posterior sampler to a computationally more demanding generic sampler and confirm that we recover the shape of the posterior. We illustrate the approach using artificial data and in an application of sign restrictions to identify oil market shocks. We show that posterior inference is sharpened compared

to the traditional approach of imposing sign restrictions and that oil supply shocks play a major role in driving oil price dynamics.

In the third chapter I investigate the effects of uncertainty shocks in the spirit of Bloom (2009) using a newly developed Bayesian Proxy Factor-augmented VAR (BP-FAVAR) model. This model combines a large information set with an identification scheme based on external instruments, thereby jointly addressing informational deficiency issues and non-credible identification assumptions. I propose a new sampling algorithm exploiting the state-space representation of the model. I find that uncertainty shocks have adverse effects on the real economy and are deflationary in the short run. To recover the dynamic causal effects, the identification scheme is crucial.

Zusammenfassung

Bayesianische Modelle sind exzellent für die Prognose und strukturelle Analyse von makroökonomischen Zusammenhängen geeignet. A-priori-Verteilungen, die in diesem Rahmen verwendet werden, erleichtern die Anwendung von Modellmittlungsverfahren, erlauben die Identifizierung von kausalen Effekten in multiplen Zeitreihenmodellen und sie ermöglichen die Schätzung von hoch-dimensionalen Modellen. Bei der Auswahl dieser a-priori-Verteilungen muss der Forscher jedoch darauf achten, dass sie seine Einschätzungen hinsichtlich der Modellparameter vor Betrachtung der Daten angemessen reflektieren. Darüber hinaus muss diese Spezifizierung eine nachvollziehbare und flexible Implementierung des Schätzverfahrens ermöglichen. In dieser Dissertation illustriere ich die Verwendung von Bayesianischen Schätzverfahren im Rahmen von Modellmittlungsverfahren, vektor-autoregressiven Modellen und einem hoch-dimensionalen Ansatz zur Identifizierung von Unsicherheitsschocks.

Im ersten Kapitel, das auf einer gemeinsamen Arbeit mit Tigran Poghosyan basiert, verwenden wir ein Frühwarnsystem (EWS), um Indikatoren zu ermitteln, die gute Prognoseeigenschaften für Fiskalkrisen haben. Insbesondere prüfen wir die Robustheit dieser Indikatoren, indem wir die Extreme Bounds Analysis (EBA) verwenden, ein Modellmittlungsverfahren, das in einer bayesianischen Version von Leamer (1985) eingeführt und von Sala-i Martin (1997) popularisiert wurde. Wir zeigen, dass sowohl fiskalische als auch nicht-fiskalische Indikatoren robuste Warnsignale für Fiskalkrisen beinhalten. In einem zweiten Schritt bewerten wir die Prognosegüte eines EWS, das auf den robustesten Frühindikatoren basiert. Wir stellen fest, dass es, verglichen mit einem Basismodell, das nur auf fiskalischen Frühindikatoren basiert, eine höhere Prognosegüte hat. Im Anschluss bewerten wir die Robustheit dieser Ergebnisse für verschiedene Modellspezifikationen, Teilproben und Schätzstrategien.

Im zweiten Kapitel, das auf einer gemeinsamen Arbeit mit Michele Piffer basiert, entwickeln wir eine neue Methode, um a-priori-Verteilungen über die Impulsantworten von strukturellen vektor-autoregressiven (SVAR) Modellen auszudrücken. Diese neue Methode unterscheidet sich von traditionellen Ansätzen dahingehend, dass die zur Verfügung stehenden a-priori-Verteilungen nicht auf eine Teilfamilie beschränkt sind,

die technisch leicht zu implementieren ist. Somit kann der Forscher flexibel seine Einschätzungen über die wahrscheinlichen Effekte von strukturellen Schocks in das Modell inkorporieren. Der Schätzalgorithmus ist aufgrund der Verwendung eines neuen importance samplers effizient und unkompliziert zu implementieren. Wir vergleichen unseren importance sampler mit einem rechenintensiveren generischen sampler und bestätigen, dass wir korrekt aus der a-posteriori Verteilung ziehen. In einem zweiten Schritt veranschaulichen wir unseren Ansatz anhand einer Anwendung von Vorzeichenrestriktionen zur Identifikation von strukturellen Schocks auf dem Ölmarkt. Wir zeigen, dass die a-posteriori-Verteilung aus unserem Ansatz gegenüber dem traditionellen Ansatz eine geringere Varianz aufweist und dass Ölangebotsschocks eine wichtige Rolle für die Dynamik des Ölpreises spielen.

Im dritten Kapitel untersuche ich die Auswirkungen von Unsicherheitsschocks im Sinne von Bloom (2009) anhand eines neu entwickelten Bayesianischen Proxy-Faktor-VAR-Modells (BP-FAVAR). Dieses Modell basiert auf einer großen Anzahl an Zeitreihen und ist somit weniger anfällig für Verzerrungen, die aus Informationsdefiziten resultieren. Gleichzeitig nutzt es externe Instrumente zur Identifikation von strukturellen Schocks und vermeidet somit schwer zu verteidigende Annahmen hinsichtlich der Reaktion von makroökonomischen Variablen auf Unsicherheitsschocks. Ich schlage einen neuen effizienten Schätzalgorithmus vor, der die Zustandsraumdarstellung des Modells nutzt. Ich stelle fest, dass Unsicherheitsschocks negative Auswirkungen auf die Realwirtschaft haben und kurzfristig deflationär wirken. Des Weiteren zeige ich, dass das Identifikationsschema von entscheidender Bedeutung für die korrekte Schätzung der dynamischen kausalen Effekte von Unsicherheitsschocks ist.

Introduction and Overview

Bayesian methods have become increasingly popular for estimating empirical macroeconomic models in recent years. The reasons for this increased popularity are two-fold: Firstly, a Bayesian set-up provides a technically convenient estimation framework. For example, it facilitates averaging parameter estimates across models and it allows to estimate high-dimensional systems. Second, the use of prior distributions allows the researcher to flexibly and transparently introduce external information about key economic parameters into the model. This feature offers an environment suitable for causal inference. These two advantages, technical convenience and the ability to introduce external information into the model, have led to a rapid increase in the number of Bayesian empirical macroeconomic applications. Still, the Bayesian methodological toolkit needs to be extended and the performance of Bayesian approaches should be investigated in different applications.

The key feature of Bayesian models, the prior distribution, can be employed for various purposes. It facilitates averaging parameter estimates across different model specifications. Model averaging techniques can be employed to assess the sensitivity of parameters, thereby yielding specification-robust results, as argued by Leamer (1985). Furthermore, within a multiple time series context, prior distributions can be employed to express prior beliefs about model parameters with a direct economic interpretation (see for example Baumeister and Hamilton, 2015). Expressing prior beliefs directly on parameters with an economic interpretation allows the researcher to flexibly and transparently incorporate prior information to achieve identification, thereby making causal inference. Lastly, prior distributions can be employed to make the estimation of high-dimensional time series models feasible by combining shrinkage of the parameter space with shrinkage of the variable space, to a degree controlled by the researcher (see Bernanke et al., 2005). A Bayesian approach allows to introduce an identification scheme based on external data to this model in a tractable and efficient manner (see Caldara and Herbst, 2019), thereby offering a framework to draw credible causal inference in a large-dimensional model.

In this thesis I contribute to the development of the Bayesian toolkit for empirical macroeconomic models and I illustrate their benefits in applications to fiscal crises, oil market shocks and uncertainty shocks. In particular, the thesis (i) uses a model averaging approach to identify robust leading indicators for fiscal distress, (ii) proposes a new approach to flexibly express prior beliefs on the impulse response functions of structural VAR models, and (iii) proposes a new Bayesian framework to combine dynamic factor analysis with an external instrument identification scheme to recover the effects of uncertainty shocks.

In the first chapter, which is based on joint work with Tigran Poghosyan, we use an Early Warning System (EWS) to recover leading indicators of fiscal distress events. In particular, we use Extreme Bounds Analysis (EBA), a model averaging approach in Bayesian spirit introduced by Leamer (1985) and popularised by Sala-i Martin (1997), to assess the robustness of leading indicators for fiscal distress across different models. We find that both fiscal and non-fiscal leading indicators are robust predictors of fiscal distress events. In a second step we assess the forecasting performance of an EWS based on the most robust leading indicators. We find that it offers a gain in predictive power compared to a baseline model which is based on fiscal leading indicators only. Lastly, we assess the robustness of these results across various model specifications, subsamples and estimation strategies.

The contribution to the literature of this chapter is three-fold: Firstly, it shows how to apply EBA in the context of a fiscal EWS and analyses its performance across various specifications. This is an important step forward compared to EWS based on an ad hoc choice of variables. Secondly, it exemplifies the importance of non-fiscal indicators for the prediction of fiscal distress events, thereby confirming theoretical results on the joint importance of sound fundamentals and resilience to speculative attacks by for example Obstfeld et al. (1984). Lastly, it offers a newly developed EWS based on both fiscal and non-fiscal leading indicators, which has a better predictive performance than a baseline model using only fiscal leading indicators.

In the second chapter, which is based on joint work with Michele Piffer, we propose a new approach to express prior beliefs on the impulse responses of structural vector auto-regressive (SVAR) models. This approach does not restrict the family of prior distributions to a subset that is technically convenient. Rather, it combines extensive flexibility in the choice of priors with an efficient importance sampler to explore the posterior distribution. We compare our new posterior sampler to a computationally more demanding generic sampler and confirm that we recover the shape of the posterior well. We illustrate the approach in an application of sign restrictions to identify oil market shocks. We show that posterior inference is sharpened compared to the

traditional approach of imposing sign restrictions and that oil supply shocks play a major role in driving oil price dynamics.

This chapter extends the Bayesian toolkit of structural VAR models and provides researchers with a methodology that allows for more flexibility in expressing prior beliefs, thereby complementing the work by Baumeister and Hamilton (2015) and Arias et al. (2018). In particular, drawing on importance sampling techniques, we show that the traditional approach to sign restrictions offers an excellent way to generate proposal draws, which can then be mapped into posterior draws from our more flexible model set-up at a low computational cost. In the empirical section, we find support for previous findings by Baumeister and Hamilton (2018) and Caldara et al. (2018) on the importance of oil supply shocks for oil price movements. We also investigate the historical role of oil supply shocks and confirm previous findings used by Antolín-Díaz and Rubio-Ramírez (2018) for identification.

In the third chapter I investigate the effects of uncertainty shocks in the spirit of Bloom (2009) using a newly developed Bayesian Proxy Factor-augmented VAR (BP-FAVAR) model. This model combines a large information set with an identification scheme based on external instruments thereby jointly addressing informational deficiency issues and non-credible identification assumptions. I propose a new sampling algorithm exploiting the state-space representation of the model. I find that uncertainty shocks have adverse effects on the real economy and are deflationary in the short run. To recover the dynamic causal effects, the identification scheme is crucial.

The chapter contributes to the literature in two ways: First, by combining a newly developed Bayesian Proxy VAR by Caldara and Herbst (2019) with a Bayesian FAVAR model, it offers a way to combine large information sets with structural identification schemes based on external information. Using a Bayesian factor approach, the researcher can directly control the degree of variable and parameter shrinkage, respectively. Second, I show that to correctly recover the effects of uncertainty shocks, the identification scheme is crucial, thereby complementing the recently growing literature investigating credible identification schemes to recover uncertainty shocks (for example Piffer and Podstawski, 2017 and Ludvigson et al., 2018).

CHAPTER 1

Leading Indicators of Fiscal Distress: Evidence from Extreme Bounds Analysis

1.1 Introduction

The global financial crisis and the subsequent weakening of fiscal positions in advanced and emerging economies once again underscored the importance of monitoring the vulnerability of countries to fiscal distress. A fiscal distress episode is a period when a government experiences extreme funding difficulties, which can manifest as outright default, debt restructuring, bond yield pressure, a large IMF-supported program, or excessive inflation. Which variables should policymakers watch to assess fiscal risks? There is a large empirical literature attempting to answer this question using Early Warning Systems (EWS); a review of this literature is in Appendix 1. International financial organizations, central banks, rating agencies, and other organizations draw this literature to develop indicators of vulnerability to fiscal distress. Most existing EWSs have the following characteristics (see Abiad, 2003 for a survey): First, they use a predetermined set of leading indicators to assess country's vulnerability to fiscal distress, which are typically based on economic reasoning. Second, the list of these variables varies widely across studies, in part driven by the preference of the researchers for parsimonious models with a large number of significant leading indicators. Finally, the results from the studies are mixed, with no agreement as to which leading indicators are most robustly associated with fiscal distress. A key characteristic of a robust leading indicator is that its coefficient's sign does not depend on the model specification.

The purpose of this paper is to revisit the issue of leading indicators of fiscal distress using Extreme Bounds Analysis (EBA). The main advantage of this methodology is

This chapter was previously published as: Bruns, Martin, and Tigran Poghosyan. "Leading indicators of fiscal distress: evidence from extreme bounds analysis." *Applied Economics* 50.13 (2018): 1454-1478.
https://doi.org/10.1080/00036846.2017.1366639.

that it takes an agnostic approach with respect to the leading indicators of fiscal distress and does not require the researcher to predetermine the set of explanatory variables. Instead, it "allows the data to speak" and ranks a set of possible leading indicators based on the "robustness" of their association with fiscal distress. This methodology is used in other economic fields, notably growth theory (see for example Sala-i-Martín, 1997) and to analyze financial crises, for example by Ho (2010). However, to our best knowledge we are the first to apply it for analyzing fiscal distress.

Our analysis leads to the following conclusions. First, both fiscal and non-fiscal leading indicators are robustly associated with fiscal distress. This is in contrast to traditional measures of fiscal stress based on fiscal leading indicators (Baldacci et al., 2011). Second, a vulnerability index based on these robust indicators performs comparably to the average performance of other EWS (for fiscal, currency, banking, and other types of crises) in the literature. However, it has better predictive power than the EWS of Baldacci et al. (2011), which is based on only fiscal leading indicators. Finally, the main result on the importance of both fiscal and non-fiscal leading indicators for fiscal distress is robust to various model specifications and sample compositions. The policy implication is that policymakers should not restrict their attention to fiscal indicators only when assessing country's vulnerability to fiscal distress.

The remainder of the paper is structured as follows. Section 1.2 describes the empirical methodology and data. Section 1.3 presents the estimation results using EBA and conducts robustness checks. Section 1.4 develops an indicator of fiscal distress based on most robust leading indicators identified using EBA and assesses its predictive power. The final section concludes.

1.2 Empirical Methodology and Data

1.2.1 Extreme Bounds Analysis

As discussed above, Extreme Bounds Analysis (EBA) is an agnostic approach to identify explanatory variables that are robustly associated with an outcome variable. Instead of pre-selecting a small set of explanatory variables, EBA explores a large number of combinations from a pool of variables used in the existing literature and based on theoretical considerations.

In the context of an Early Warning System (EWS) for fiscal distress events, the following logit specification is estimated:

$$log(\frac{P(Y=1|z,x_j)}{P(Y=0|z,x_j)}) = \beta_{zj}z + \beta_{xj}x_j + \epsilon \tag{1.1}$$

Where Y is a binary outcome variable indicating a fiscal distress event, z is the leading indicator whose robustness is being assessed and x_j are other leading indicators from the pool of all leading indicators, excluding z. Assume there are n elements in this pool. Then, for each leading indicator z, one has to estimate a regression with k additional controls x_j, which gives $M = (n-1-k)!/k!$ combinations per leading indicator z and a total of nM regressions to be estimated.

For each leading indicator z, we follow the following procedure:

- Estimate each of the M regressions and store the estimated coefficient of z, $\hat{\beta}_{zj}$.
- Weigh coefficients $\hat{\beta}_{zj}$ by the relative likelihood of the model that they are a part of, so that coefficients stemming from a very unlikely model receive less weight than others. The weight is:

$$\omega_{zj} = \frac{L_{zj}}{\sum_{i=1}^{M} L_{zj}} \tag{1.2}$$

- Calculate the weighted average of coefficients across regressions to obtain the coefficient of leading indicator z:

$$\hat{\beta}_z = \sum_{j=1}^{M} \omega_{zj}\beta_{zj} \tag{1.3}$$

- Calculate a likelihood-weighted average standard error in a similar fashion.
- Assess the robustness of $\hat{\beta}_z$. Assuming that $\hat{\beta}_z$ is normally distributed across regressions, calculate the proportion of its distribution to one side of zero, and consider $\hat{\beta}_z$ to be robust if this proportion exceeds a certain confidence level (e.g. 90 percent).

Following these steps for each indicator results in a set of indicators that can be ranked according to their respective robustness.

1.2.2 Data

Our dataset covers 29 advanced and 52 emerging economies (81 countries in total) over the 1970 to 2015 period. It builds on the original dataset of fiscal distress developed by Baldacci et al. (2011), extending the data through 2015 and expanding the set of leading indicators.

1.2.2.1 Fiscal distress

A fiscal distress episode is identified as a period when government experiences extreme funding difficulties. Based on the literature survey, Baldacci et al. (2011) identify four types of criteria to capture fiscal distress episodes: (i) debt default or restructuring; (ii) sovereign bond yield pressure; (iii) large IMF-supported program; and (iv) excessive inflation. We adopt the same definition of fiscal distress (see Table 1) and expand the series through 2015.

Table 1.1: Definition of Fiscal Distress

Event	Criteria	Advanced Economies	Emerging Economies
Public debt default or restructuring	Failure to service debt as payments come due, as well as distressed debt exchanges	S&P definition	S&P definition
Extreme financing constraint of the sovereign	Sovereign yield pressure	Sovereign spreads greater than 1,000 basis points or 2 s.d. from the country average	Sovereign spreads greater than 1,000 basis points or 2 s.d. from the country average
Large financing	Large IMF-supported program	Access to 100 percent of quota or more	Access to 100 percent of quota or more
Implicit/Internal public debt default	High inflation rate	Inflation greater than 35 percent per annum	Inflation greater than 500 percent per annum

Note: Baldacci et al. (2011)

Table 1.2 presents the distribution of fiscal distress episodes across their types and country groups. Several observations emerge. First, comparison across country groups suggests that advanced economies have experienced fewer fiscal distress episodes than emerging economies. The low unconditional likelihood of fiscal distress for advanced economies will have implications for the conditional analysis conducted below. Second, the comparison across types of distress events suggests that advanced economies experienced almost no outright default episodes and most fiscal distress episodes manifested in the form of bond yield pressures. By contrast, emerging economies experienced a large number of fiscal distress episodes and were frequent recipients of large IMF-supported bailouts. Finally, the total number of distress episodes is lower than the

sum of individual episodes. This is due to the fact that some countries experienced multiple fiscal distress events simultaneously.

Table 1.2: Fiscal Distress Episodes

	Sovereign Default or Restructuring	Bond Yield Pressures	Inflation Pressures	IMF program	Total
Advanced Economies	1 (0)	30 (29)	5 (5)	8 (6)	41 (39)
Emerging Economies	54 (52)	17 (15)	6 (6)	103 (79)	160 (135)

Note: Baldacci et al. (2011) and author's calculation for 2011-2015. The numbers in brackets represent fiscal distress episodes in the original Baldacci et al. (2011) dataset running through 2010.

Figure 1.1 shows the share of fiscal distress episodes that coincided with currency and banking crises using the Reinhart and Rogoff dataset.[1] We find that currency crises overlap with 26 percent of fiscal distress episodes, while banking crises overlap with 24 percent of fiscal distress episodes for the sample spanning through 2010 (the last year of Reinhart and Rogoff's dataset). The high share of overlaps suggest that some of the fiscal distress events may have originated outside of the fiscal sector, which is consistent with the "twin crises" narrative proposed by Kaminsky and Reinhart (1999).

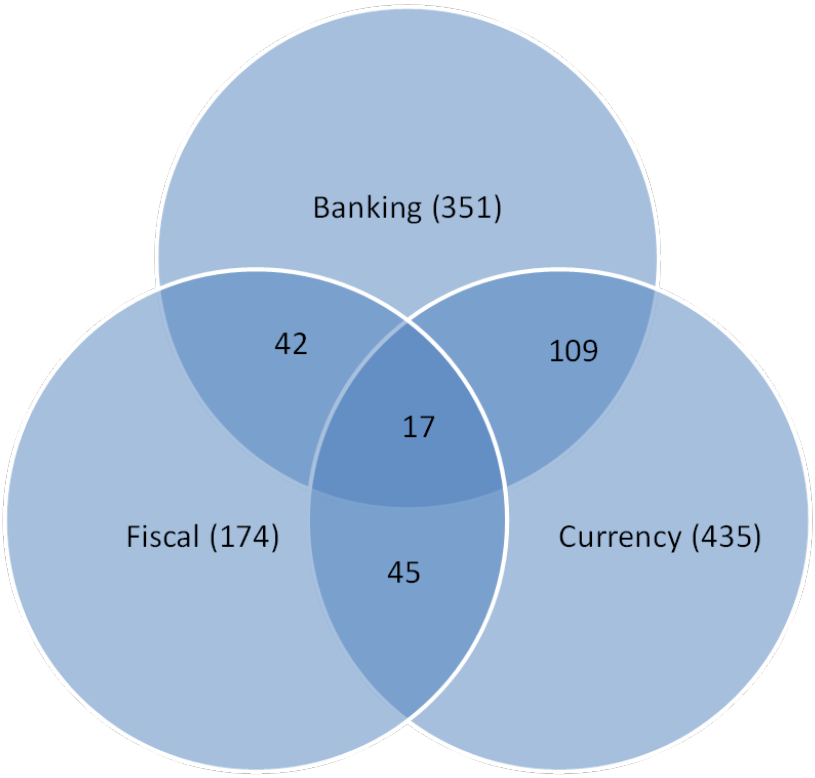
1.2.2.2 Leading Indicators

We identify 37 variables that were used as leading indicators for fiscal distress in the literature. Table 1.6 in Appendix 1.A.2 lists these variables as well as their sources. The indicators cover not only the fiscal sector, but also macro, monetary, and external sectors. Some of the indicators reflect the impact of contagion effects and global factors, which is consistent with predictions of the theoretical literature. All variables are measured in percentage points. The panel is unbalanced as some indicators are not available for all countries and differ in time coverage.

Table 1.7 presents descriptive statistics, while Table 1.8 presents bivariate correlations based on the pooled series. As shown in the latter tables, most variables are not highly correlated with the exception of some debt-related indicators. This should limit multicollinearity issues.

[1]The dataset is available at: http://www.reinhartandrogoff.com/data/.

Figure 1.1: Banking, Currency and Fiscal Crises



Note: Reported is the overlap between fiscal, banking, and currency crises for the period 1970–2010. Data on banking and currency crisis episodes are from Reinhart and Rogoff (2009). Data on the first year of a fiscal crisis are from Baldacci et al. (2011).

1.3 Estimation Results

In this section, we present estimation results for the baseline specification covering the whole sample and based on the pooled logit specification. We also check whether the results are sensitive to changes in the sample and estimation methods. Leading indicators are assessed to be robust if the rate of confidence in the sign of the coefficient exceeds 90 percent, as approximated by the normal distribution. We use a forecasting horizon of 1 year throughout the estimation.[2]

1.3.1 Baseline Extreme Bounds Analysis Model

Table 1.3 presents the baseline EBA estimation results. Estimations are performed using the pooled logit model. We use various combinations of two additional controls per regression, which leads to 630 regressions per leading indicator (23,410 regressions in total). The main reason for using two additional controls as a baseline is computa-

[2]The results do not change qualitatively when using a forecasting horizon of two years.

tional efficiency and a higher number of observations per regression. The variables are ranked according to their robustness, with the most robust indicators placed on top.

The estimation results suggest that both fiscal and non-fiscal variables are robust leading indicators of fiscal distress. We find that fiscal distress tends to follow a period of overheating in the real sector (widening of output gap). This is consistent with the Alberola et al. (2013) finding that growth, if it is not driven by sound economic fundamentals, can be detrimental to systemic stability. A related explanation could be that a large positive output gap, even if it is non-inflationary, can be associated with a build-up of financial imbalances, which in turn can indirectly lead to fiscal distress (see Borio et al., 2016).

In addition, adverse developments in the external sector (high current account deficit, low level of FX reserves/GDP ratio, slowdown in FX reserves growth, and higher openness) tend to precede fiscal distress episodes. These results are consistent with both Chakrabarti and Zeaiter (2014)'s analysis of external debt arrears as well as Gourinchas and Obstfeld (2012) analysis of the causes of sovereign default. A negative effect of openness is found in Manasse et al. (2003).

Among fiscal leading indicators, we found that foreign exchange debt to GDP ratio, primary fiscal balance gap, as well as primary and overall fiscal balance to GDP ratios, are robust indicators, which is in line with most fiscal EWSs, such as Baldacci et al. (2011).

One potential reason for the robustness of non-fiscal indicators could be a high correlation with fiscal indicators. Table 1.8 shows that this is not the case: for example, the correlation between FX reserves (% of GDP) and various robust fiscal indicators, such as primary balance, overall balance or primary balance gap does not exceed 0.35 in absolute terms. The same is true for the output gap. This indicates that non-fiscal indicators do not just pick up the effects of fiscal indicators, but that they matter on their own.

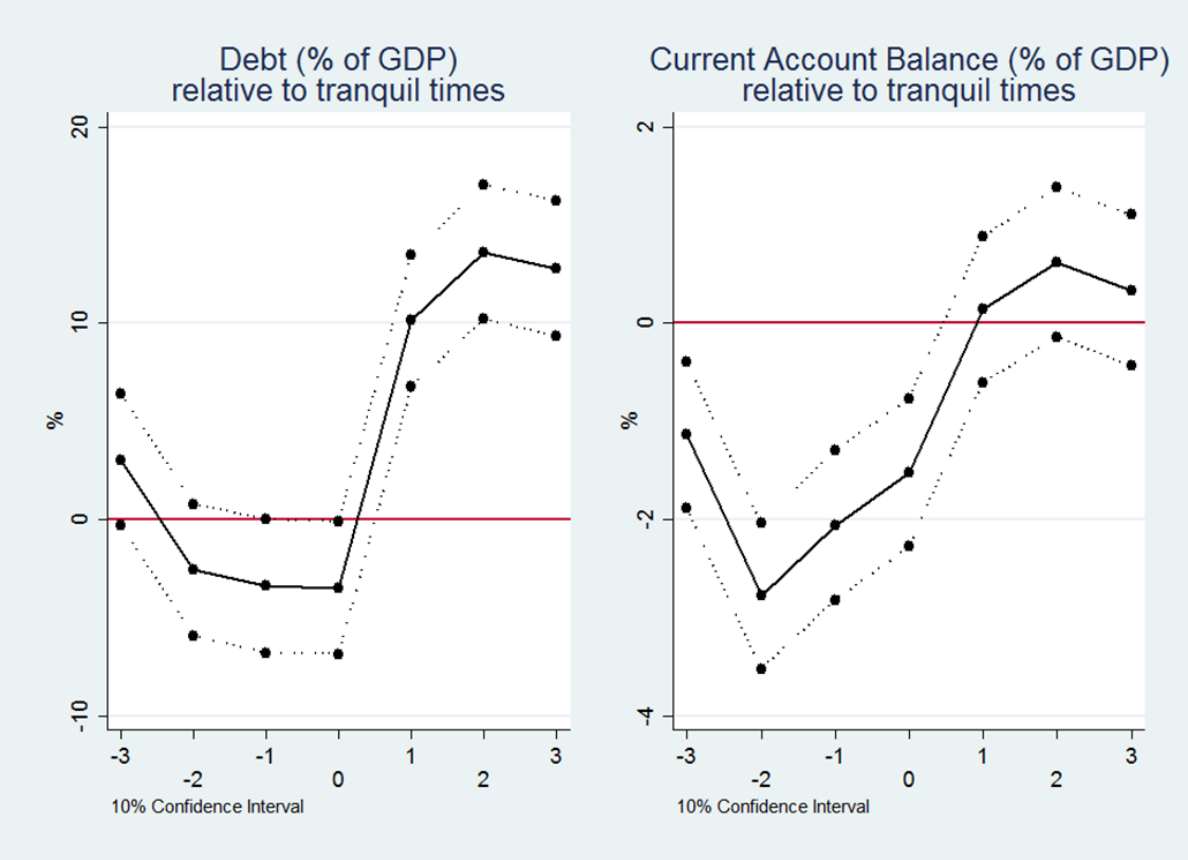
An illustration for why external factors perform better than fiscal factors as leading indicators is observable in Figure 1.2. It shows an application of the Event Study Methodology proposed by Gourinchas and Obstfeld (2012), where a fixed effects model is employed to regress the leading indicator of interest on dummies indicating the distance from a fiscal distress event. The graphs plot the estimated coefficients of these dummies, which are a measure of the percent deviation of the respective leading indicator from its "tranquil" time average. This "tranquil" time average is defined as the average level of the respective variable outside the Event Study horizon. Figure 1.2 shows that the debt to GDP ratio before the crisis is not significantly different from levels observed in tranquil times, but increases rapidly after the crisis incident. This

Table 1.3: Extreme Bounds Analysis (baseline)

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.296	0.069	1.000	3188
Current Account Balance (% of GDP)	-0.046	0.014	1.000	3275
FX Reserves Growth	-0.010	0.003	0.998	3051
FX Reserves (% of GDP)	-0.034	0.012	0.998	3103
Openness: (M+X)/ GDP	-0.006	0.003	0.970	3254
Primary Balance Gap (% of GDP)	-0.029	0.017	0.958	688
Real GDP Growth	-0.039	0.023	0.957	3269
Overall Fiscal Balance (% of GDP)	-0.052	0.031	0.953	2035
Primary Balance (% of GDP)	-0.048	0.037	0.907	1874
Foreign Exchange Debt (% of GDP)	0.008	0.006	0.903	1124
Gross Financing Need	0.016	0.013	0.885	735
Change in Net Claims on Central Government	0.000	0.000	0.861	2415
Short Term Debt (% of total)	-0.008	0.007	0.854	2107
Unemployment Rate	-0.020	0.021	0.826	2546
LIBOR	0.042	0.049	0.801	3670
FX Reserve Coverage	-0.012	0.015	0.788	1581
CPI Inflation	0.004	0.005	0.786	2804
Concessional Debt (% of total)	-0.004	0.006	0.766	1632
Interest Expenditure (% of total Expenditure)	0.010	0.014	0.763	1908
Domestic Credit Gap	0.001	0.002	0.755	2588
Real Interest Rate	0.010	0.015	0.752	1387
Short Term External Debt (% of GDP)	-0.005	0.008	0.746	2142
Debt (% of GDP)	-0.002	0.004	0.744	3073
Real Exchange Rate Undervaluation	-0.002	0.004	0.741	2877
External Debt (% of GDP)	-0.002	0.003	0.722	2475
Average Maturity	-0.020	0.039	0.698	733
FDI (% of GDP)	0.010	0.033	0.625	2972
Amortisation of Total Public Debt (% of GDP)	-0.003	0.008	0.622	2030
Debt owed to Commercial Banks (% of total)	-0.005	0.016	0.618	1423
US TBill Rate	0.010	0.042	0.594	3592
GDP per Capita	-0.006	0.033	0.575	3291
Short Term Debt (% of FX Reserves)	0.000	0.001	0.568	1984
Nominal GDP Growth	0.001	0.005	0.567	3046
Debt Service due (% of GDP)	-0.003	0.020	0.551	1602
Debt owed to Multilateral Creditors (% of total)	0.001	0.009	0.545	1632
Interest-Growth Differential	-0.001	0.008	0.527	2154
Average Effective Interest Rate	0.000	0.002	0.523	1951

Note: Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the pooled logit model, with 2 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015.

Figure 1.2: Event Study Analysis



Note: Both graphs as generated using the Gourinchas and Obstfeld (2012) Event Study Methodology. It shows the percent deviation of the Debt Ratio and the Current Account Balance from their respective "tranquile" average. This is defined as the average level of the respective variable outside the window of (-3; 3) around the crisis incident. Dotted lines denote the 10% Confidence Interval.

indicates that the debt to GDP ratio - one of the key fiscal indicators used in previous studies (see, e.g., Kraay and Nehru, 2006) - is more of an ex post indicator of fiscal distress rather than a leading indicator. The current account balance, on the other hand, is significantly lower relative to tranquil times before the crisis, suggesting that it can be used as a leading indicator. This result supports the hypothesis that fiscal vulnerabilities can be built up outside the public sector.

Overall, the baseline results suggest that limiting leading indicators to fiscal variables, like in Baldacci et al. (2011), may result in a loss of important information regarding vulnerability to fiscal distress.[3] Consistent with predictions of most recent theories of crises, information from other sectors should also be monitored.

[3]Our results remain qualitatively unchanged when the sample is restricted to the period 1970–2010 as in Baldacci et al. (2011).

1.3.2 Robustness Checks

We run several additional EBA regressions to check the results for sensitivity to changes in the sample and estimation methods (see Appendix 1.A.3).

"Narrow" definition of fiscal distress The robustness of non-fiscal leading indicators, especially those pertaining to the external sector, may be driven by the fact that some of the large scale IMF programs (the third definition of fiscal distress) are triggered by balance of payment or other crises that are not purely linked to fiscal distress. In addition, high inflationary pressures (fourth definition of fiscal distress) may be caused by large devaluations following currency crises. To check whether the robustness of non-fiscal variables still holds for "narrow" fiscal distress episodes, we re-estimate the model using a more restrictive sample of fiscal distress episodes that did not overlap with banking and currency crises.

Table 1.9 in Appendix 1.A.3 presents estimation results from a specification that restricts the dependent variable to "narrow" fiscal distress episodes. The set of robust determinants still includes non-fiscal leading indicators, confirming the baseline results. Gross financing needs and interest costs also turn robust in this specification, suggesting that borrowing cost pressures are important leading indicators for "narrow" fiscal distress episodes. The main drawback of this specification is that the sample has to be restricted through 2010, the last year for which banking and currency crisis variables are available in the Reinhart and Rogoff database.

Emerging economies As indicated in the previous section, most fiscal crisis events took place in emerging economies, which could explain the relatively high proportion of external leading indicators that are mostly relevant for emerging economies. For example, difficulties to finance a persistent current account deficit and the associated decline in FX reserves have historically been associated with fiscal distress in emerging markets. Advanced economies on the other hand, face fewer financing problems of this sort as they are often able to borrow in their own currency and generally rely less on external financing (IMF, 2010). To assess whether our baseline results are mostly driven by emerging economies, we redo the analysis by restricting the sample to emerging economies.[4]

Table 1.10 in Appendix 1.A.3 presents estimation results from the sample of emerging economies. The set of robust indicators is very similar to the one for the total sample, confirming the influence of emerging economies in driving the main results. Out of 10 robust leading indicators found in the total sample, eight remain significant in emerging economies. One additional variable that becomes significant for emerging economies is

[4]Unfortunately, we could not replicate the analysis for just the advanced economies due to the limited number of fiscal distress events.

the unemployment rate. The latter is consistent with the robust result on the output gap, given that unemployment is its mirror image.

Number of control variables We also check the robustness of results to the larger number of control variables. Table 1.11 in Appendix 1.A.3 presents the estimation results from a specification that increases the number of control variables from two to three. This results in a substantially increased time for model estimation, as the number of regressions per leading indicator increases to 7140 (264,180 regressions in total). Nevertheless, the main results remain unchanged, with 8 out of 10 robust indicators from the baseline regression remaining robust. As before, the set of robust leading indicators includes both fiscal and non-fiscal variables.

Random effects logit and pooled probit models We also check the robustness of the main results to types of discrete choice models. First, we control for unobserved country-specific heterogeneity by using a random effects logit model.[5] Table 1.12 in Appendix 1.A.3 presents the estimation results, showing that the set of robust indicators remains qualitatively unchanged, with 8 out of 10 robust indicators from baseline regression remaining robust. Second, given that the logit model has fatter tails than the probit model, we also check the sensitivity to tail risks by using a pooled probit model. Table 1.13 in Appendix 1.A.3 presents estimation results, showing that the set of robust indicators remains practically unchanged and suggesting that fat tails are not affecting the results.

Using alternative filtering technique for the output gap Throughout the analysis, we find a persistent evidence of robustness of the output gap variable, which comes on top of the list in all tables. This could be due to the fact that the two-sided HP filter used to estimate the output gap could potentially bias the results prior to fiscal crises as it uses future values of output, already affected by the crisis, to smooth past series. To check the sensitivity of the results to the smoothing technique, we use the one-sided filter of Christiano and Fitzgerald (2003). Table 1.14 in Appendix 1.A.3 presents estimation results using this filter. The results remain unchanged suggesting that the measurement of the output gap does not constitute a problem and reinforcing the importance of output cycles as leading indicators of fiscal crises.

[5]We have also tried a fixed effects logit specification and the results remain similar. The main drawback of the fixed effect model is that it drops countries that have never experienced a fiscal distress (see Bussiere, 2013 for a discussion).

1.4 Fiscal Distress Index and Its Performance

In this section, we present an alternative fiscal distress index based on the most robust leading indicators identified using the EBA methodology (Table 1.4).[6] We then measure its performance for in-sample and out-of-sample and make a comparison with other EWS in the literature.

A natural question is which indicators make the largest contribution to the fiscal distress index. In order to assess this, Figure 1.3 shows the impact of changes in robust leading indicators between the 25th and the 75th percentile on the logarithm of the odds ratio of the fiscal crisis. The log odds ratio is given in equation (1.1). It is based on the results of the Logit model presented in Table 1.4. In order to compute it, we take a specific variable and compute its 25th and 75th quantile. Then we take the difference these two quantiles and multiply it by the corresponding coefficient $\hat{\beta}_i$. The result writes as $(\hat{\beta}_i x_i^{(0.75)} - \hat{\beta}_i x_i^{(0.25)})$, where $x_i^{(q)}$ is the q-th quantile of leading indicator i. It measures by how much the log odds ratio changes if indicator i moves from its 25th to its 75th quantile while the remaining indicators stay constant.[7]

The figure shows that non-fiscal indicators have a larger relative impact on the fiscal distress index than fiscal indicators reflecting their importance in assessing underlying vulnerabilities.

1.4.1 Fiscal Distress Index

Figure 1.4 presents the evolution of the fiscal distress index based on most robust leading indicators for advanced and emerging economies and the fiscal distress incidents. This index is obtained by computing the predicted values from the logit model, which includes the most robust leading indicators obtained in the first step. Two observations are worth noting. First, the level of fiscal stress tends to be lower in advanced economies than in emerging economies, which is consistent with the fact that advanced economies have a lower propensity to experience distress on average. Second, there are spikes in the level of fiscal distress around periods of notable financial crises (late 1990s, the global financial crisis). This is consistent with the empirical regularity of clustering across different types of crises due to contagion effects.

[6] The index does not include primary balance gap variable (because of low numbers of observations), the foreign exchange debt ratio (because the series end in 2012), and the FX reserves to GDP ratio (because we already control for the FX reserves growth variable).

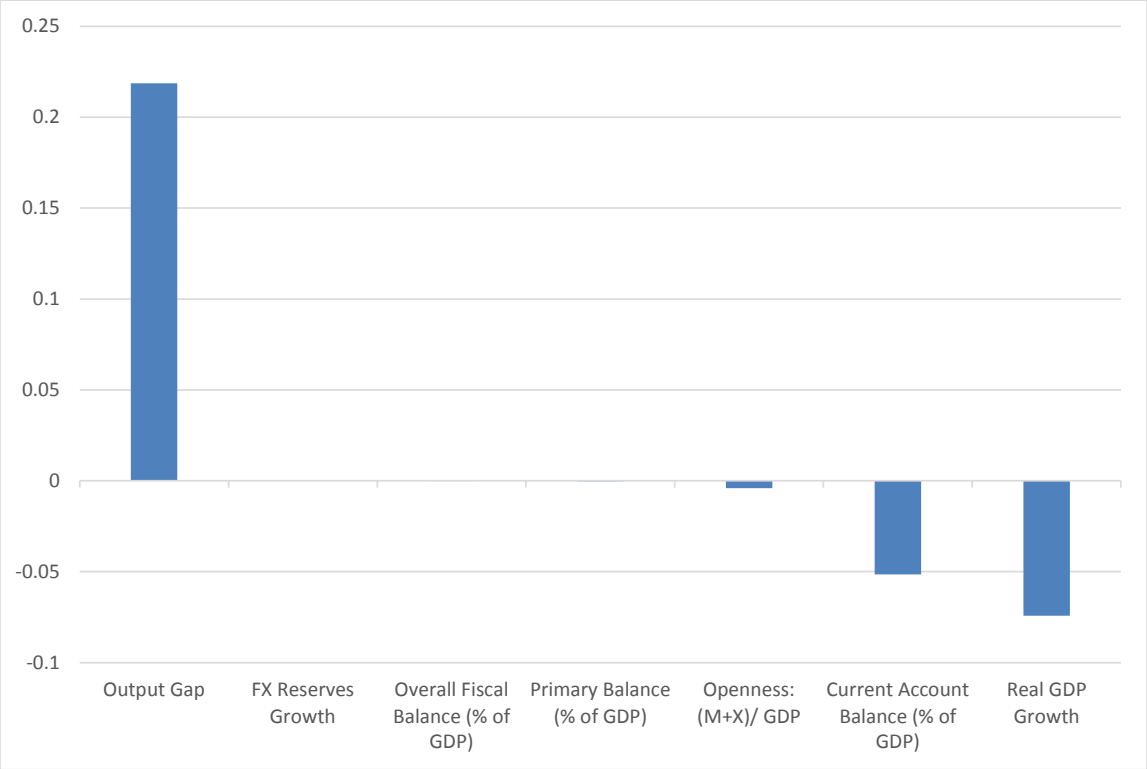
[7] As an example, consider the Current Account Balance (% of GDP). Its 25th quantile is -0.0998 and its 75th quantile is 0.4465. The coefficient, as given in Table 1.4, is $\hat{\beta}_i = -0.0943$. This gives a log odds ratio of $(0.4464 + 0.0998)(-0.0943) = -0.0515$ as shown in Figure 1.3. This means that when Current Account Balance (% of GDP) moves from its 25th to its 75th quantile, holding the other variables constant, the log odds ratio will change by -0.0515.

Table 1.4: Early warning system based on the most robust leading indicators

	Pooled logit
Output Gap	0.4377***
	[0.0860]
Current Account Balance (% of GDP)	-0.0943***
	[0.0196]
FX Reserves Growth	0.0002
	[0.0029]
Openness: (M+X)/ GDP	-0.0097***
	[0.0035]
Real GDP Growth	-0.1312***
	[0.0349]
Overall Fiscal Balance (% of GDP)	-0.0030
	[0.0407]
Primary Balance (% of GDP)	-0.0078
	[0.0439]
Intercept	-1.9855***
	[0.3069]
Number of observations	1765
Pseudo R^2	0.537
AUROC	0.815
Type 1 error	0.218
Type 2 error	0.312
TME	0.530

Note: Reported are estimation results from the pooled logit model using most robust leading indicators of fiscal distress. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015. ***, **, and * denote significance at 1, 5, and 10 percent levels, respectively.

Figure 1.3: Economic Significance of Robust Leading Indicators



Note: Reported is the difference in the logarithm of the odds ratio calculated for 75th and 25th percentile of each leading indicator (components of the fiscal distress index). A larger absolute value indicates a larger economic impact of a change in that indicator on the log odds ratio of a fiscal distress.

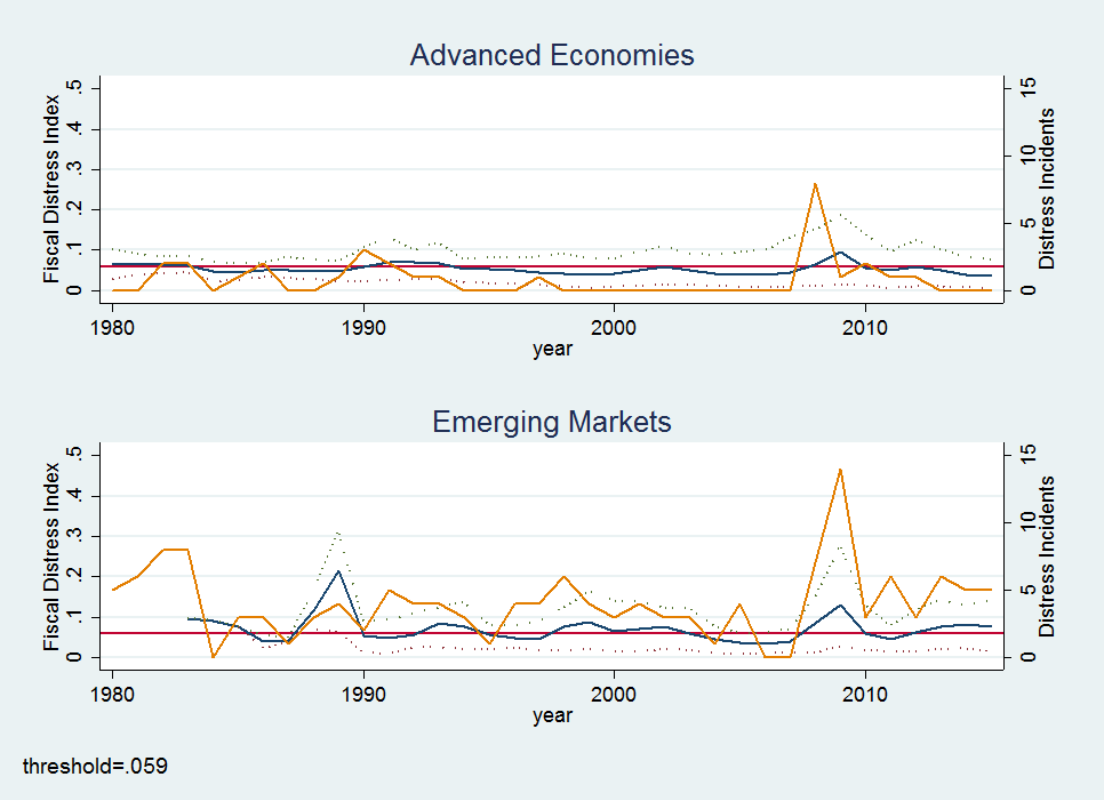
1.4.2 Measure of Performance

As pointed out by Ho (2015), it is not obvious that the most robust leading indicators are also those that yield the best predictive performance of an EWS. This is why, following Baldacci et al. (2011), we assess the performance of our model using the total misclassification error:

$$TME(t) = Type1(t) + Type2(t), \tag{1.4}$$

where $Type1(t)$ indicates missed crises as a share of all crises, $Type2(t)$ indicates false alarms as a share of all non-crisis periods, and t is the threshold level. The threshold can take values between 0 and 1, thus dividing the fiscal distress index resulting from the Logit model into two regions: When the index exceeds the threshold level, an alarm is raised. When the index stays below the threshold, an alarm is not raised. The choice of the threshold level is determined by the policymaker facing a trade-off.

Figure 1.4: Fiscal Distress Index Based on Robust Leading Indicators

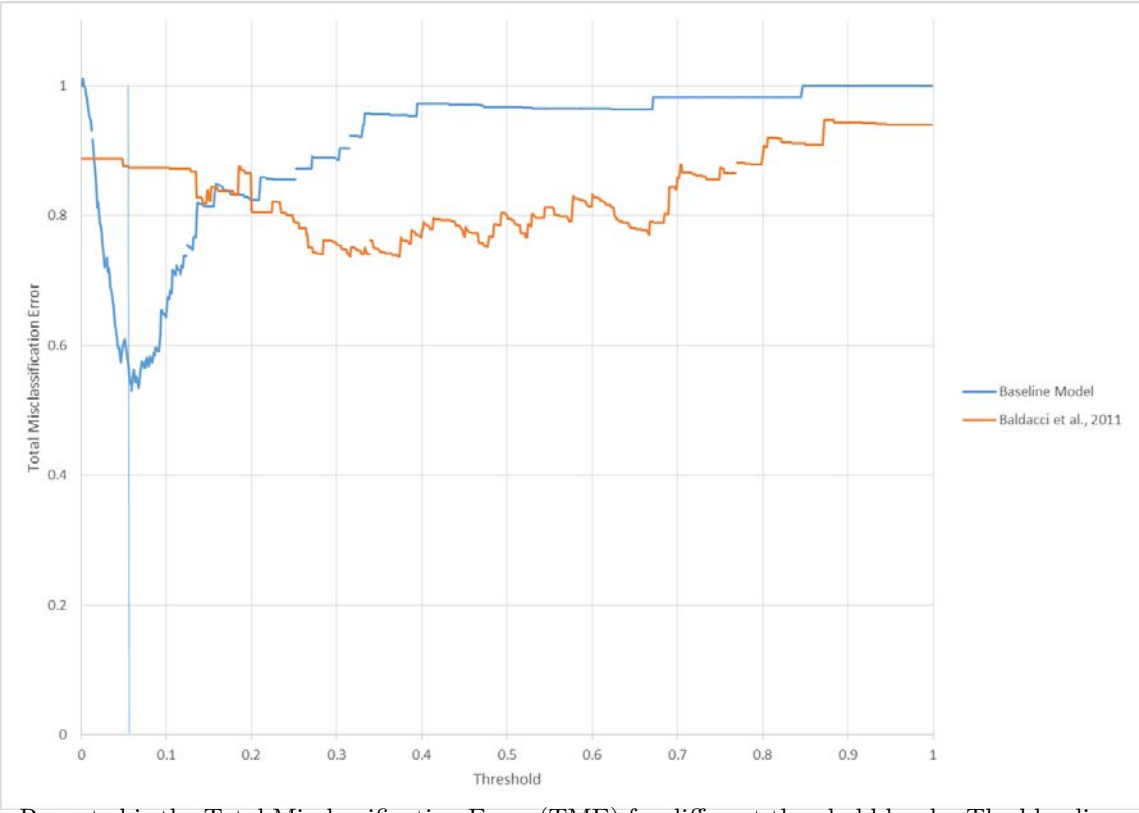


Note: The fiscal distress index is generated by using predicted values from the logit model that includes most robust leading indicators. The most robust leading indicators are defined as those with at least 90 percent probability of not switching signs using the EBA results. Robust leading indicators include (i) Output Gap; (ii) Current Account Balance (% of GDP); (iii) FX Reserves Growth; (iv) Openness: (M+X)/GDP; (v) Real GDP Growth, (vi) Overall Fiscal Balance (% of GDP); and (vii) Primary Balance (% of GDP). The blue line indicates the Fiscal Distress Index (left axis). Dotted lines indicate 90 percent confidence bands around predicted values. Orange line represents the absolute number of distress events in a given year (right axis). Data are missing for the pre-1980 period, meaning that no predictive values can be generated.

If t is set too high, the index will cross it in very few cases and many crises will be missed, thus resulting in a large Type 1 error. Alternatively, if t is set too low, the index will frequently cross the threshold, meaning that many alarms will be falsely issued resulting in a large Type 2 error. We assume that the utility of the policy maker can be represented by using TME as a loss function, which is minimized over t.

Figure 1.5 presents the in-sample performance for different threshold values for both the EBA-based indicator and the Baldacci et al. (2011) fiscal distress index. Neither model is dominated by the other, as the performance depends on the choice of the threshold. Figure 1.5 shows that the TME of the EBA-based indicator is minimized at 0.52 when the level of the threshold is 0.07. The Baldacci et al. (2011) fiscal distress index obtains a minimum of 0.73 when the level of the threshold is 0.38. Thus, the

Figure 1.5: In-sample Performance: EBA-based Fiscal Distress Index versus Baldacci et al. (2011) Index

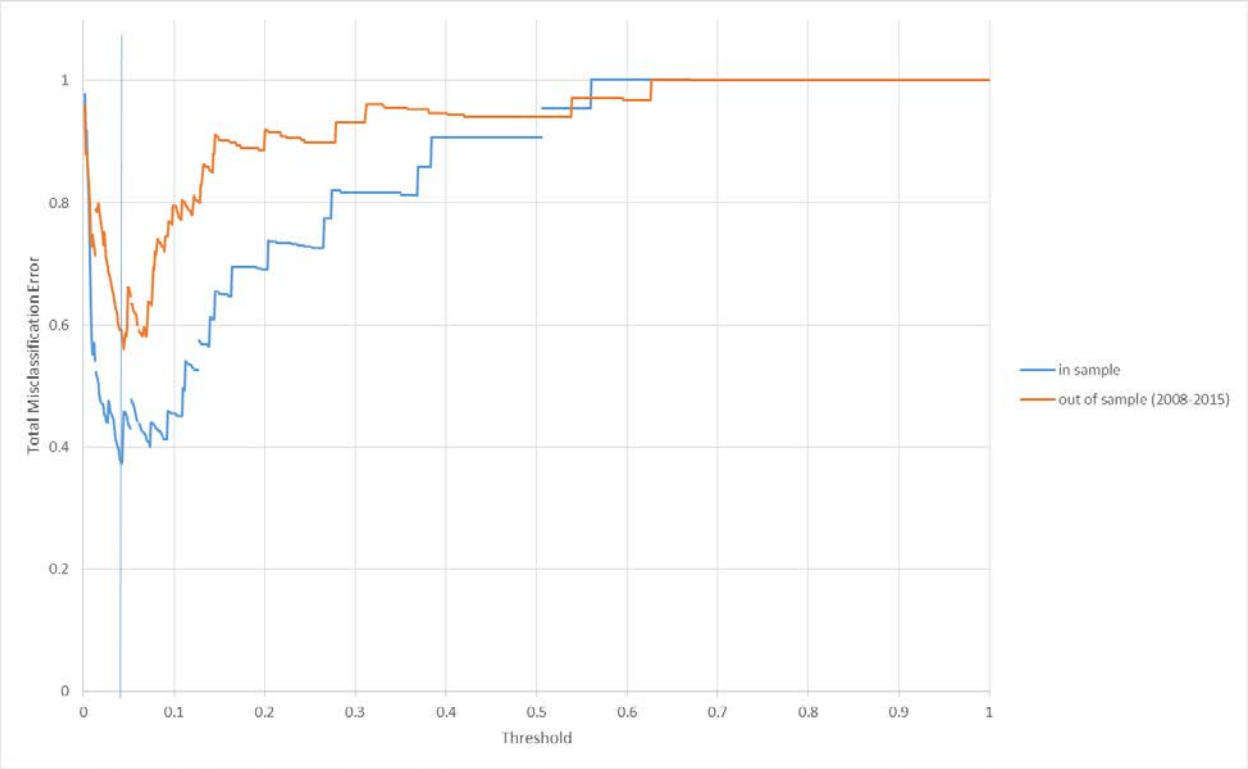


Note: Reported is the Total Misclassification Error (TME) for different threshold levels. The blue line represents the baseline model based on robust leading indicators: (i) output gap; (ii) current account balance (% of GDP); (iii) FX reserves growth; (iv) openness: (M+X)/GDP; (v) real GDP growth; (vi) overall fiscal balance (% of GDP); and (vii) primary balance (% of GDP). The vertical line marks the threshold level (0.07) that minimizes the TME. The orange line represents TME from the model based on variables used by Baldacci et al. (2011).

TME is reduced by 29 percent for the optimal choice of the threshold, suggesting that our model has a better fit.

Figure 1.6 presents in-sample and out-of-sample performance for different threshold values of the EBA-based fiscal distress indicator. We use the years 1970 through 2007 to fit the model, compute the coefficients and obtain the threshold. The model is then applied to the years 2008 through 2015, computing the fiscal distress index by combining these estimated coefficients from the model-fitting sample with the data in the prediction sample. As expected, the in-sample performance outperforms out-of-sample performance. Nevertheless, the minimum TME for the out-of-sample is not much higher than in-sample TME suggesting that our model would have done well predicting the post-2007 sample using information available through 2007.

Figure 1.6: In-sample Versus Out-of-sample Performance of the EBA-based Fiscal Distress Index



Note: Reported is The Total Misclassification Error (TME) for different levels of the threshold. The fiscal vulnerability index is based on robust leading indicators: (i) output gap; (ii) current account balance (% of GDP), (iii) FX reserves growth; (iv) openness: (M+X)/GDP; (v) real GDP growth; (vi) overall fiscal balance (% of GDP); (vii) primary balance (% of GDP). The blue line represents the in-sample TME based on the years 1970-2007. The orange line represents the out-of-sample TME for the years 2008-2015. The vertical line marks the threshold level (0.04) that minimizes the in-sample TME.

Table 1.5 presents the classification table for out-of-sample predictions made by the EBA-based fiscal distress indicator. It is an illustration for the two types of errors that every EWS makes, as explained above. Out of the 32 crises that were identified between 2008 and 2015, the model predicts 25 correctly (78 %). A false alarm is issued for 118 out of 344 non-crisis years (34%).

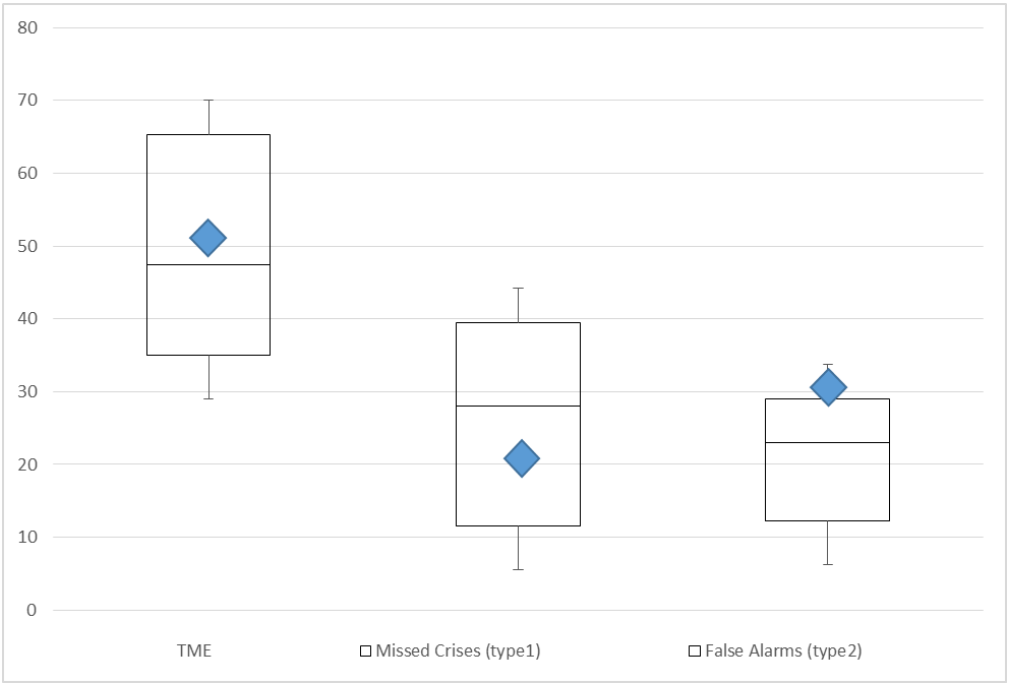
Figure 1.7 presents a comparison with other studies. It plots the minimum, 25th percentile, mean, 75th percentile and maximum of the 3 relevant quality measures of an EWS: The TME, type 1 error, and type 2 error It shows that our model is within the 1st and 3rd quartile of the TME when compared to other studies on EWS. It performs slightly worse than average regarding the type 2 errors and slightly better regarding the type 1 errors. As mentioned in Berg et al. (2005), a different trade-off between the two error types could be achieved using an asymmetrically weighted loss function.

Table 1.5: Classification Table for Out-of-sample Predictions of the EBA-based Fiscal Distress Indicator

		Crisis		
		Yes	No	Total
Signal	Yes	25	118	143
	No	7	226	233
	Total	32	344	376

Note: Reported is the model classification for the years 2008–2015. An alarm is issued if fiscal distress index exceeds the threshold derived from the in-sample minimization of the TME.

Figure 1.7: Comparison with Other Studies



Note: Reported is the in-sample total misclassification error (TME), Type 1 and Type 2 errors of eight Early Warning Systems for currency, banking and sovereign crises reported in the literature. The values represent the minimum, the 25th percentile, the mean, the 75th percentile and the maximum.

1.5 Conclusion

Reviewing the large literature on EWS for fiscal distress, we assess the robustness of the leading indicators employed in these studies using Extreme Bounds Analysis. We find both fiscal and non-fiscal leading indicators to be robust.

In a second step, we build a vulnerability index using the most robust leading indicators and find that its predictive properties are close to the average found in other EWS

(for fiscal, currency, banking, and other types of crises) in the literature. However, the use of robust fiscal and non-fiscal leading indicators helps us to improve performance relative to Baldacci et al. (2011) EWS based on fiscal indicators only. From a policy perspective, these results suggest that some non-fiscal leading indicators should be monitored as closely as fiscal leading indicators to assess a country's vulnerability to fiscal distress.

Designing a EWS based on robust leading indicators poses several problems that could be addressed in the future. First, the estimation sample using a logit model is constrained by the shortest time series, which can significantly reduce the sample size and, thus, precision. Second, while using a pooled sample of both advanced and emerging economies does increase the sample size, it constrains the leading indicators and their relative effects to be the same across the two subsamples. However, separate analysis for advanced economies only is complicated given that most fiscal crises have occurred in emerging economies. Lastly, identifying "narrow" fiscal distress events is challenging due to the frequent occurrence of twin or multiple crises and the spillovers among them.

1.A Further results

1.A.1 Related Literature

1.A.1.1 Theories of Crisis Determinants

When modeling crises, it is important to draw distinction between underlying vulnerabilities and crisis risks. The presence of underlying vulnerabilities is a necessary precondition for a crisis to occur, but not sufficient. Crises tend to be triggered by external or domestic shocks, which are highly uncertain. The underlying vulnerability helps to identify countries that are prone to crisis should even a moderate shock occur.

The early theoretical literature emphasizes the role of fundamentals in measuring underlying vulnerability. The classic reference is Krugman (1979), which predicts that weak fundamentals, in part driven by unsustainable fiscal and monetary policies, make countries vulnerable to a balance of payments crisis. For fiscal crises, Detragiache and Spilimbergo (2001) show that sovereign default is the only equilibrium response to a large negative shock to fundamentals.

The following generation of theoretical literature of crises emphasizes the role of self-fulfilling expectations and non-fiscal fundamentals. The self-fulfilling crisis literature (Obstfeld et al., 1984; Calvo, 1988; Alesina et al., 1989; Cole and Kehoe (1996); Jeanne, 1997; and Masson, 1999) was inspired by the fact that while some crises were preceded by deterioration in fundamentals, some speculative attacks have taken place without apparent monetary and fiscal imbalances. These studies developed multiple-equilibrium models that better allow capturing the complex interaction between underlying vulnerabilities and speculative attacks (Eichengreen et al., 1995 provide a review). In these models, countries can jump from good to bad equilibrium for a certain range of values of economic fundamentals. For example, Calvo (1988) shows that if a government is unable to commit to repay its debt, multiple equilibria, including repudiation and inflation, can arise.

The main takeaway from the review of theoretical literature is that there is no reason a priori to restrict leading indicators of fiscal distress to fiscal fundamentals.

1.A.1.2 Empirical Evidence on Early Warning Indicators of Fiscal Distress

The empirical literature on early warning indicators of fiscal distress was developed as part of a larger literature on early warning indicators of financial crises. Comprehensive surveys of EWS for banking, currency, sovereign debt, equity, and inflation crises can be found in Kaminsky and Reinhart (1999), Klau et al. (2000), d Abiad et al. (2003), Berg et al. (2005), and Frankel and Saravelos (2012). As discussed above, the main

objective of these studies is to identify leading indicators making countries vulnerable to a crisis, rather than prediction of the timing of the crisis. The empirical techniques used in this literature could be grouped into two main categories: non-parametric and parametric (Baldacci et al., 2011; Frankel and Saravelos, 2012; and Comelli, 2013).

The most popular non-parametric EWS is the "signals" approach popularized byKaminsky et al. (1998) in the context of currency crises. This approach selects a number of variables as leading indicators of crises and determines threshold values beyond which a crisis signal is considered to have been given. The main drawback of this approach is that it only focuses on bivariate association between an early warning indicator and crises, thus controlling for other factors is not allowed. In addition, the statistical significance of the early warning indicators cannot be determined directly, although out-of-sample performance could be assessed. Studies using these techniques in the context of fiscal crises include Reinhart (2002), Baldacci et al. (2011), Berti et al. (2012) and De Cos et al. (2014). These studies suggest that fiscal fundamentals, such as the level of public debt, the composition of public debt, fiscal deficit, and fiscal financing needs are important leading indicators of fiscal distress.

The parametric EWS models draw on limited dependent variable techniques (multivariate logit, probit). These methods allow testing the significance of various leading indicators in determining the likelihood of crisis occurring in the near future, while accounting for their correlation. However, these methods require long-time series of leading indicators and low degrees of freedom that may prevent the use of multivariate approach when the number of predictors is large. Studies applying parametric methods to analyze fiscal crises include Marashaden (1997), Peter (2002), Manasse et al. (2003), Kraay and Nehru (2006), and Gourinchas and Obstfeld (2012). Sumner et al. (2017) apply a logit approach to European fiscal stress events, updating the Baldacci et al. (2011) series of fiscal stress events. Their in-sample model performance is comparable to ours, while the out-of-sample performance cannot be directly compared due to differences in sample splits. These studies confirm the importance of fiscal fundamentals, but suggest that macroeconomic developments in general, especially in the external sector, also play a role.

Both groups of studies share a common characteristic – the set of leading indicators of fiscal distress is predetermined by the researcher. This selection is typically done with the benefit of hindsight, with the significance of the leading indicator typically playing a role in the selection process. To overcome this issue, some studies apply extreme bounds analysis to study leading indicators of financial crises.

1.A.1.3 Extreme Bounds Analysis

Extreme bounds analysis was developed and applied to study determinants of growth by Leamer (1985), and was later on extended and popularized by Sala-i Martin (1997). The methodology does not rely on a predetermined set of explanatory variables and "lets the data speak" by examining all possible combinations of explanatory variables. The focus is on the change in signs of explanatory variables. If a sign change is observed relatively frequently, it is said that the explanatory variable is not robustly related to the dependent variable.

Early warning studies using extreme bounds analysis (or its Bayesian equivalent) include Chakrabarti and Zeaiter (2014), Ho (2015), and Christofides et al. (2016). However, to the best of our knowledge no study used this approach to analyze determinants of fiscal distress, which is the gap we fill with our paper. The closest paper to ours is Chakrabarti and Zeaiter (2014), which analyzes determinants of external debt arrears using a linear regression model. The results suggest that a range of fiscal and non-fiscal indicators, including growth, inflation, trade deficit, foreign reserves, and exchange rates, are robust predictors of external debt arrears.

Alessi et al. (2015) provide an overview of alternative ways of combining the information in a large set of variables, such as the Lasso, the ridge regression estimator, Bayesian Model Averaging, principal component analysis, and factor models. However, an advantage of Extreme Bounds Analysis is that it does not introduce an estimation bias, unlike the Lasso or the ridge regression. Furthermore, its data requirements are less strict than for Bayesian Model Averaging as it can be applied to unbalanced panels and the results are more directly interpretable than those stemming from principal component or factor models.

1.A.2 Descriptive Statistics

Table 1.6: List of leading indicators of fiscal distress

Leading Indicator	Source	Grouping	Description
Average Maturity	Baldacci et al. (2011)	Public Sector: Liquidity Pressure Indicators	Average Maturity of remaining debt (in years)
Gross Financing Need	Baldacci et al. (2011)	Public Sector: Debt Burden	Gross financing needs (short-term debt plus the overall balance), percent of GDP
Interest-Growth Differential	Baldacci et al. (2011)	Public Sector: Debt Tolerance Indicators	Difference between average effective interest rate and GDP growth in percent
Primary Balance (% of GDP)	Baldacci et al. (2011)	Public Sector: Liquidity Pressure Indicators	Revenue-Expenditure+Interest Expenditure, percent of GDP
Short Term Debt (% of FX Reserves)	Baldacci et al. (2011)	Public Sector: Debt Burden	Short Term Debt, percent of FX Reserves
Short Term Debt (% of total)	Baldacci et al. (2011)	Public Sector: Debt Burden	Short Term Debt, percent of total Debt
GDP per Capita	Chakrabarti and Zeaiter (2014)	Public Sector: Debt Tolerance Indicators	Gross domestic product per capita, current prices in U.S. dollars
Concessional Debt (% of total)	Detragiache and Spilimbergo (2001)	Public Sector: Debt Burden	Concessional Debt, percent of total Debt
Debt owed to Commercial Banks (% of total)	Detragiache and Spilimbergo (2001)	Public Sector: Debt Burden	Debt owed to Commercial Banks, percent of total Debt
Debt owed to Multilateral Creditors (% of total)	Detragiache and Spilimbergo (2001)	Public Sector: Debt Burden	Debt owed to Multilateral Creditors, percent of total Debt
Debt Service due (% of GDP)	Detragiache and Spilimbergo (2001)	Public Sector: Liquidity Pressure Indicators	Debt Service due, percent of GDP
Current Account Balance (% of GDP)	Gourinchas and Obstfeld (2012)	Public Sector: Liquidity Pressure Indicators	Balance on current account, percent of GDP
Domestic Credit Gap	Gourinchas and Obstfeld (2012)	Public Sector: Macroeconomic Factors	Domestic Credit percent deviation from HP-filter trend
FX Reserves (% of GDP)	Gourinchas and Obstfeld (2012)	Public Sector: Macroeconomic Factors	Foreign Exchange Reserves, percent of GDP
Real Exchange Rate Undervaluation	Gourinchas and Obstfeld (2012)	Public Sector: Macroeconomic Factors	Percent deviation of real effective exchange rate from HP filter trend
Real Interest Rate	Gourinchas and Obstfeld (2012)	Public Sector: Macroeconomic Factors	6 months Treasury Bill Rate - CPI Inflation (end of period percentage change)
Short Term External Debt (% of GDP)	Gourinchas and Obstfeld (2012)	Public Sector: Debt Burden	Short Term External Debt, percent of GDP
Change in Net Claims on Central Government	Hemming et al. (2003)	Public Sector: Liquidity Pressure Indicators	Change in Net Claims on Central Government, in percent
CPI Inflation	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Consumer Prices, end-of-period, percent change
Debt (% of GDP)	Manasse et al. (2003)	Public Sector: Debt Burden	Public Debt, percent of GDP
External Debt (% of GDP)	Manasse et al. (2003)	Public Sector: Debt Burden	Total external debt, gross, including arrears and other short-term debt, percent of GDP
FDI (% of GDP)	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Foreign direct investment, net inflows in reporting economy, percent of GDP
FX Reserves Growth	Manasse et al. (2003)	Public Sector: Debt Tolerance Indicators	Annual percentage change in Foreign Exchange Reserves
Interest Expenditure (% of total Expenditure)	Manasse et al. (2003)	Public Sector: Liquidity Pressure Indicators	Interest Expenditure, percent of total Expenditure
LIBOR	Manasse et al. (2003)	Contagion Effects and Global Factors	London Interbank Offered Rate
Nominal GDP Growth	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Nominal GDP, annual percentage change
Openness: (M+X)/ GDP	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Exports+Imports, percent of GDP
Overall Fiscal Balance (% of GDP)	Manasse et al. (2003)	Public Sector: Liquidity Pressure Indicators	Revenue-Expenditure, percent of GDP
Real GDP Growth	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Gross domestic product, constant prices, annual percentage change
Unemployment Rate	Manasse et al. (2003)	Public Sector: Macroeconomic Factors	Unemployment Rate
US TBill Rate	Manasse et al. (2003)	Contagion Effects and Global Factors	6 months US Treasury Bill Rate
Output Gap	Ostry et al. (2010)	Public Sector: Macroeconomic Factors	Percent deviation of real GDP from Baxter-King filter
Average Effective Interest Rate	VEE 2014	Public Sector: Liquidity Pressure Indicators	Interest payments(t)/General Government Debt(t-1)
Foreign Exchange Debt (% of GDP)	VEE 2014	Public Sector: Debt Burden	General government gross debt in foreign currency, percent of GDP
FX Reserve Coverage	VEE 2014	External Sector	Reserves, percent of (Short Term debt at remaning maturity+Current Account deficit)
Primary Balance Gap (% of GDP)	VEE 2014	Public Sector: Liquidity Pressure Indicators	Primary Balance Gap, percent of GDP
Amortisation of Total Public Debt (% of GDP)	VEE 2015	External Sector	Amortization paid, (principal only), percent of GDP

Source: Survey of the literature by the authors.

Table 1.7: Descriptive statistics

	Obs.	**Mean**	**Med.**	**St. Dev.**	**10th percentile**	**90th percentile**
Debt (% of GDP)	3,073	52.5	45.3	34.1	16.6	95.2
Output Gap	3,188	-0.1	-0.1	2.0	-1.5	1.3
Openness: (M+X)/ GDP	3,254	71.2	60.0	53.2	26.1	123.0
CPI Inflation	2,804	28.3	6.2	224.0	1.4	26.1
FX Reserve Coverage	1,581	1.7	0.7	11.2	0.1	2.9
Amortisation of Total Public Debt (% of GDP)	2,030	10.0	4.2	19.9	0.6	19.4
Real Interest Rate	1,387	-1.2	1.4	52.9	-4.9	6.9
Domestic Credit Gap	2,588	-48.1	-1.4	1754.8	-94.0	26.3
Current Account Balance (% of GDP)	3,275	-2.1	-1.9	7.6	-8.4	4.7
Real Exchange Rate Undervaluation	2,877	-3.6	-1.5	38.0	-24.6	20.3
FX Reserves (% of GDP)	3,103	12.3	7.5	15.0	1.9	25.9
Short Term External Debt (% of GDP)	2,142	14.4	4.8	37.8	0.0	27.6
Interest-Growth Differential	2,154	-15.7	-4.1	108.6	-24.7	3.0
Primary Balance (% of GDP)	1,874	0.7	0.3	3.9	-3.4	5.3
Short Term Debt (% of total)	2,107	17.9	13.8	17.5	0.0	41.3
Short Term Debt (% of FX Reserves)	1,984	654.0	51.6	5328.2	0.0	438.5
Debt owed to Commercial Banks (% of total)	1,423	11.3	5.8	14.0	0.1	30.6
Concessional Debt (% of total)	1,632	22.8	15.6	22.1	1.3	58.9
Debt owed to Multilateral Creditors (% of total)	1,632	17.4	14.1	13.1	3.7	35.3
Debt Service due (% of GDP)	1,602	5.8	4.3	6.4	1.4	10.9
FX Reserves Growth	3,051	18.3	9.4	52.4	-19.8	58.6
LIBOR	3,670	2.5	2.6	2.5	-1.1	5.4
US TBill Rate	3,592	5.1	5.0	3.3	0.1	9.6
Unemployment Rate	2,546	8.5	7.2	6.0	2.6	15.7
Nominal GDP Growth	3,046	40.4	11.2	297.2	2.8	34.9
Real GDP Growth	3,269	3.5	3.7	4.9	-0.9	8.3
External Debt (% of GDP)	2,475	68.2	44.4	90.0	14.5	141.1
FDI (% of GDP)	2,972	-1.4	-0.7	3.4	-5.1	1.0
Overall Fiscal Balance (% of GDP)	2,035	-2.5	-2.5	4.3	-7.4	2.1
Interest Expenditure (% of total Expenditure)	1,908	9.9	7.8	8.9	2.3	20.2
Change in Net Claims on Central Government	2,415	47.4	10.8	795.9	-48.3	116.4
GDP per Capita	3,291	9.3	9.7	4.2	6.0	13.3
Average Effective Interest Rate	1,951	17.5	0.2	75.1	0.0	14.4
Gross Financing Need	735	11.4	8.8	11.3	1.3	26.6
Average Maturity	733	7.2	6.2	4.2	3.0	13.3
Primary Balance Gap (% of GDP)	688	0.9	0.8	8.3	-4.7	7.6
Foreign Exchange Debt (% of GDP)	1,124	15.4	4.4	22.6	0.0	41.5

Source: IMF WEO, World Development Indicators, International Finance Statistics, Government Finance Statistics.

Table 1.8: Correlation Matrix

	Debt (% of GDP)	Output Gap	Openness: (M+X)/ GDP	CPI Inflation	FX Reserve Coverage	Amortisation of Total Public Debt (% of GDP)	Real Interest Rate	Domestic Credit Gap	Current Account Balance (% of GDP)	Real Exchange Rate Undervaluation	FX Reserves (% of GDP)	Short Term External Debt (% of GDP)
Debt (% of GDP)	1.000											
Output Gap	-0.173	1.000										
Openness: (M+X)/ GDP	-0.160	-0.007	1.000									
CPI Inflation	0.249	-0.070	-0.034	1.000								
FX Reserve Coverage	-0.272	-0.057	0.291	-0.199	1.000							
Amortisation of Total Public Debt (% of GDP)	-0.421	0.063	0.106	0.017	-0.260	1.000						
Real Interest Rate	0.353	-0.054	-0.650	0.064	-0.326	-0.108	1.000					
Domestic Credit Gap	0.059	0.299	-0.143	0.002	-0.551	0.108	0.063	1.000				
Current Account Balance (% of GDP)	-0.451	-0.087	-0.238	-0.323	0.378	0.403	0.160	-0.156	1.000			
Real Exchange Rate Undervaluation	0.202	0.012	0.290	0.096	-0.017	0.195	-0.247	0.062	-0.083	1.000		
FX Reserves (% of GDP)	-0.223	0.114	0.715	-0.151	0.282	0.397	-0.562	-0.156	0.096	0.150	1.000	
Short Term External Debt (% of GDP)	-0.595	0.110	0.026	-0.162	-0.224	0.906	-0.138	0.223	0.462	0.064	0.302	1.000
Interest-Growth Differential	0.556	0.108	-0.414	-0.384	-0.238	-0.149	0.346	0.062	-0.044	0.026	-0.168	-0.154
Primary Balance (% of GDP)	0.652	0.042	-0.374	0.106	-0.232	-0.281	0.491	-0.152	-0.218	-0.177	-0.269	-0.385
Short Term Debt (% of total)	-0.665	0.295	0.039	-0.184	-0.184	0.703	-0.173	0.252	0.364	-0.117	0.299	0.868
Short Term Debt (% of FX Reserves)	-0.546	0.085	-0.323	-0.003	-0.361	0.642	0.087	0.245	0.297	-0.047	-0.226	0.804
Debt owed to Commercial Banks (% of total)	-0.059	-0.011	-0.092	0.563	-0.417	0.357	0.343	0.011	-0.114	-0.295	-0.095	0.295
Concessional Debt (% of total)	-0.215	-0.222	0.604	-0.028	0.498	-0.325	-0.489	-0.173	-0.102	0.158	0.215	-0.332
Debt owed to Multilateral Creditors (% of total)	0.265	-0.276	0.307	0.012	0.306	-0.672	-0.139	-0.067	-0.300	0.098	-0.068	-0.712
Debt Service due (% of GDP)	0.444	-0.200	-0.226	0.093	-0.329	0.485	0.326	-0.047	0.148	0.267	0.043	0.271
FX Reserves Growth	-0.075	-0.027	-0.163	0.004	0.347	-0.205	0.013	-0.325	0.252	-0.361	0.159	-0.124
LIBOR	-0.194	0.190	0.024	-0.261	0.193	-0.186	-0.033	-0.132	-0.011	-0.321	0.087	-0.060
US TBill Rate	-0.190	0.022	-0.090	-0.339	0.271	-0.151	0.026	-0.347	0.123	-0.358	0.055	-0.045
Unemployment Rate	0.219	0.035	-0.456	-0.003	-0.371	-0.177	0.311	0.083	-0.216	0.092	-0.578	-0.057
Nominal GDP Growth	-0.020	-0.132	-0.047	0.872	-0.075	0.037	0.129	-0.094	-0.129	-0.030	-0.185	-0.094
Real GDP Growth	-0.680	0.010	-0.070	-0.031	0.267	0.233	-0.055	-0.118	0.437	-0.262	-0.047	0.327
External Debt (% of GDP)	0.355	-0.256	-0.007	0.084	-0.216	0.358	0.053	0.045	-0.017	0.493	0.005	0.202
FDI (% of GDP)	-0.265	-0.105	-0.629	-0.164	0.069	0.329	0.324	-0.047	0.710	-0.158	-0.273	0.455
Overall Fiscal Balance (% of GDP)	-0.596	0.283	0.281	-0.664	0.306	0.168	-0.440	-0.156	0.268	-0.168	0.349	0.326
Interest Expenditure (% of total Expenditure)	0.878	-0.183	-0.390	0.316	-0.337	-0.294	0.547	0.006	-0.293	0.064	-0.354	-0.461
Change in Net Claims on Central Government	0.159	0.047	0.178	0.133	0.102	-0.059	-0.018	0.078	-0.030	0.083	0.148	-0.145
GDP per Capita	-0.088	-0.096	0.849	-0.037	0.416	-0.314	-0.601	-0.129	-0.320	0.250	0.382	-0.339
Average Effective Interest Rate	-0.307	-0.141	0.642	0.098	0.333	-0.237	-0.461	-0.080	-0.194	0.081	0.194	-0.236
Gross Financing Need	0.585	-0.341	-0.617	0.421	-0.313	-0.014	0.693	-0.053	0.095	0.061	-0.450	-0.190
Average Maturity	0.254	0.089	-0.061	-0.260	0.218	-0.332	-0.185	-0.026	-0.172	0.321	-0.084	-0.324
Primary Balance Gap (% of GDP)	-0.138	-0.093	-0.108	0.262	0.242	-0.021	-0.182	-0.258	0.194	-0.251	0.028	-0.048
Foreign Exchange Debt (% of GDP)	0.842	-0.207	0.162	0.311	-0.148	-0.351	0.086	-0.008	-0.575	0.343	-0.075	-0.552

Interest-Growth Differential	Primary Balance (% of GDP)	Short Term Debt (% of total)	Short Term Debt (% of FX Reserves)	Debt owed to Commercial Banks (% of total)	Concessional Debt (% of total)	Debt owed to Multilateral Creditors (% of total)	Debt Service due (% of GDP)	FX Reserves Growth	LIBOR	US TBill Rate	Unemployment Rate	Nominal GDP Growth	Real GDP Growth	External Debt (% of GDP)	FDI (% of GDP)	Overall Fiscal Balance (% of GDP)	Interest Expenditure (% of total Expenditure)
1.000																	
0.673	1.000																
-0.131	-0.285	1.000															
-0.160	-0.249	0.733	1.000														
-0.297	0.206	0.359	0.427	1.000													
-0.615	-0.593	-0.399	-0.429	-0.417	1.000												
-0.241	-0.156	-0.741	-0.689	-0.470	0.818	1.000											
0.444	0.370	-0.066	0.108	0.153	-0.493	-0.378	1.000										
0.078	0.146	-0.007	-0.128	0.009	-0.077	-0.040	-0.149	1.000									
0.149	0.003	0.041	-0.147	-0.162	0.044	0.056	-0.248	0.342	1.000								
0.221	0.151	0.014	-0.098	-0.135	-0.060	-0.039	-0.092	0.462	0.876	1.000							
0.494	0.392	0.015	0.350	0.037	-0.418	-0.170	0.035	-0.006	0.085	0.081	1.000						
-0.617	-0.056	-0.121	0.121	0.590	0.101	0.033	-0.061	0.054	-0.304	-0.317	-0.102	1.000					
-0.645	-0.476	0.281	0.441	0.195	0.229	-0.121	-0.268	0.126	-0.172	-0.023	-0.316	0.377	1.000				
0.194	0.012	-0.259	0.121	-0.117	-0.052	-0.028	0.742	-0.284	-0.226	-0.162	0.117	-0.023	-0.182	1.000			
0.167	-0.035	0.358	0.553	0.066	-0.442	-0.511	0.249	0.169	-0.104	0.060	0.225	-0.050	0.344	0.102	1.000		
0.053	-0.115	0.491	0.163	-0.268	0.054	-0.210	-0.327	0.148	0.348	0.414	-0.073	-0.538	0.211	-0.384	0.031	1.000	
0.571	0.790	-0.568	-0.325	0.113	-0.385	0.115	0.561	-0.019	-0.142	-0.068	0.379	0.094	-0.559	0.421	-0.049	-0.609	1.000
-0.162	-0.139	-0.256	-0.299	-0.173	0.300	0.316	0.037	-0.073	-0.152	-0.291	-0.404	0.068	-0.063	0.110	-0.354	-0.235	0.046
-0.462	-0.423	-0.292	-0.478	-0.335	0.841	0.655	-0.508	-0.143	0.012	-0.126	-0.341	0.006	-0.013	-0.109	-0.673	0.162	-0.358
-0.715	-0.592	-0.204	-0.231	-0.144	0.825	0.567	-0.606	-0.097	-0.088	-0.233	-0.392	0.276	0.321	-0.221	-0.513	0.074	-0.523
0.276	0.442	-0.417	-0.019	0.245	-0.384	-0.041	0.629	0.041	-0.252	-0.138	0.219	0.340	-0.170	0.460	0.319	-0.757	0.761
0.289	-0.045	-0.439	-0.331	-0.682	0.131	0.200	0.038	-0.082	0.101	0.106	-0.024	-0.284	-0.180	0.276	-0.105	-0.025	0.055
-0.198	0.200	0.011	0.057	0.164	-0.080	-0.155	-0.051	0.333	-0.317	-0.107	-0.169	0.390	0.363	-0.172	0.141	0.105	0.006
0.286	0.439	-0.717	-0.530	-0.098	0.109	0.447	0.382	-0.213	-0.168	-0.173	0.164	0.077	-0.566	0.561	-0.457	-0.560	0.771

Change in Net Claims on Central Government	GDP per Capita	Average Effective Interest Rate	Gross Financing Need	Average Maturity	Primary Balance Gap (% of GDP)	Foreign Exchange Debt (% of GDP)
1.000						
0.237	1.000					
0.247	0.866	1.000				
0.116	-0.579	-0.478	1.000			
0.070	0.141	0.018	-0.048	1.000		
-0.024	-0.065	0.066	0.046	-0.158	1.000	
0.185	0.232	-0.048	0.410	0.243	-0.132	1.000

(continued)

1.A.3 Robustness Checks

Table 1.9: Robustness check 1: "Narrow" fiscal distress episodes

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.194	0.076	0.995	2955
FX Reserves Growth	-0.011	0.006	0.979	2736
CPI Inflation	0.025	0.012	0.978	2576
FX Reserves (% of GDP)	-0.049	0.024	0.977	2788
Primary Balance Gap (% of GDP)	-0.105	0.057	0.968	453
Gross Financing Need	0.032	0.017	0.965	502
Real Interest Rate	0.064	0.037	0.959	1263
Openness: (M+X)/ GDP	-0.010	0.006	0.951	2859
Current Account Balance (% of GDP)	-0.037	0.024	0.940	2882
Change in Net Claims on Central Government	0.000	0.000	0.939	2415
Overall Fiscal Balance (% of GDP)	-0.083	0.054	0.939	1640
Interest Expenditure (% of total Expenditure)	0.042	0.030	0.920	1513
FX Reserve Coverage	-0.023	0.021	0.870	1385
Average Maturity	-0.064	0.062	0.850	514
Nominal GDP Growth	0.012	0.012	0.832	2812
External Debt (% of GDP)	-0.005	0.006	0.831	2081
Debt owed to Multilateral Creditors (% of total)	-0.012	0.015	0.783	1509
Foreign Exchange Debt (% of GDP)	0.009	0.011	0.782	994
Concessional Debt (% of total)	-0.006	0.010	0.730	1509
Debt Service due (% of GDP)	-0.030	0.055	0.708	1479
Short Term External Debt (% of GDP)	-0.011	0.020	0.702	1834
Real Exchange Rate Undervaluation	-0.003	0.008	0.671	2646
Average Effective Interest Rate	0.003	0.008	0.651	1556
Interest-Growth Differential	-0.006	0.017	0.649	1920
Short Term Debt (% of total)	-0.005	0.013	0.645	1797
GDP per Capita	0.018	0.057	0.627	2896
Debt owed to Commercial Banks (% of total)	0.006	0.018	0.620	1423
Domestic Credit Gap	0.001	0.002	0.609	2588
US TBill Rate	0.015	0.060	0.600	3280
Debt (% of GDP)	-0.001	0.006	0.558	2678
Short Term Debt (% of FX Reserves)	0.000	0.001	0.551	1736
LIBOR	0.009	0.073	0.547	3280
Unemployment Rate	-0.004	0.039	0.544	2175
Amortisation of Total Public Debt (% of GDP)	0.002	0.021	0.541	1767
Real GDP Growth	-0.004	0.045	0.533	2874
Primary Balance (% of GDP)	-0.004	0.061	0.528	1479
FDI (% of GDP)	0.003	0.067	0.516	2579

Note: Reported are estimation results from the EBA regression. The dependent variable is the narrower fiscal stress definition, which excludes Currency and Banking crises as identified in Reinhart and Rogoff (2008). Estimations are performed using the pooled logit model, with 2 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2010.

Table 1.10: Robustness check 2: Emerging economies

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.259	0.070	1.000	3188
FX Reserves Growth	-0.012	0.004	0.999	3051
FX Reserves (% of GDP)	-0.040	0.013	0.999	3103
Current Account Balance (% of GDP)	-0.033	0.014	0.989	3275
Real GDP Growth	-0.050	0.023	0.987	3269
Overall Fiscal Balance (% of GDP)	-0.066	0.035	0.972	2035
Primary Balance Gap (% of GDP)	-0.029	0.017	0.958	688
Unemployment Rate	-0.035	0.023	0.937	2546
Primary Balance (% of GDP)	-0.056	0.042	0.911	1874
Openness: (M+X)/ GDP	-0.004	0.003	0.896	3254
Gross Financing Need	0.016	0.013	0.885	735
Change in Net Claims on Central Government	0.000	0.000	0.865	2415
FDI (% of GDP)	0.037	0.041	0.816	2972
FX Reserve Coverage	-0.012	0.015	0.784	1581
Short Term Debt (% of total)	-0.006	0.008	0.770	2107
Concessional Debt (% of total)	-0.004	0.006	0.766	1632
LIBOR	0.039	0.054	0.766	3670
CPI Inflation	0.004	0.006	0.750	2804
Short Term External Debt (% of GDP)	-0.007	0.010	0.749	2142
Domestic Credit Gap	0.001	0.002	0.748	2588
External Debt (% of GDP)	-0.002	0.004	0.719	2475
Average Maturity	-0.020	0.039	0.698	733
Real Interest Rate	0.008	0.015	0.696	1387
Real Exchange Rate Undervaluation	-0.002	0.005	0.658	2877
Interest-Growth Differential	0.003	0.009	0.638	2154
Foreign Exchange Debt (% of GDP)	0.002	0.007	0.629	1124
Short Term Debt (% of FX Reserves)	0.000	0.001	0.622	1984
Debt (% of GDP)	-0.001	0.004	0.621	3073
Debt owed to Commercial Banks (% of total)	-0.005	0.016	0.618	1423
Average Effective Interest Rate	-0.001	0.003	0.607	1951
US TBill Rate	0.006	0.047	0.552	3592
GDP per Capita	0.004	0.035	0.551	3291
Debt Service due (% of GDP)	-0.003	0.020	0.551	1602
Nominal GDP Growth	-0.001	0.006	0.547	3046
Debt owed to Multilateral Creditors (% of total)	0.001	0.009	0.545	1632
Amortisation of Total Public Debt (% of GDP)	-0.001	0.009	0.542	2030
Interest Expenditure (% of total Expenditure)	0.001	0.014	0.531	1908

Note: Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the pooled logit model, with 2 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 52 emerging economies for the period 1970-2015.

Table 1.11: Robustness check 3: Number of control variables

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.333	0.083	1.000	3188
Current Account Balance (% of GDP)	-0.054	0.017	0.999	3275
FX Reserves (% of GDP)	-0.038	0.014	0.996	3103
FX Reserves Growth	-0.011	0.004	0.993	3051
Real GDP Growth	-0.045	0.028	0.947	3269
Openness: (M+X)/ GDP	-0.006	0.004	0.941	3254
Primary Balance Gap (% of GDP)	-0.028	0.019	0.935	688
Overall Fiscal Balance (% of GDP)	-0.050	0.038	0.907	2035
Unemployment Rate	-0.030	0.025	0.891	2546
CPI Inflation	0.009	0.008	0.858	2804
Primary Balance (% of GDP)	-0.045	0.043	0.853	1874
Short Term Debt (% of total)	-0.009	0.009	0.849	2107
Gross Financing Need	0.014	0.014	0.842	735
Domestic Credit Gap	0.003	0.003	0.834	2588
Change in Net Claims on Central Government	0.000	0.000	0.786	2415
LIBOR	0.054	0.069	0.785	3670
Concessional Debt (% of total)	-0.005	0.007	0.766	1632
FX Reserve Coverage	-0.011	0.016	0.761	1581
Debt (% of GDP)	-0.003	0.005	0.755	3073
Short Term External Debt (% of GDP)	-0.008	0.011	0.754	2142
Real Interest Rate	0.015	0.022	0.753	1387
External Debt (% of GDP)	-0.003	0.004	0.746	2475
Foreign Exchange Debt (% of GDP)	0.005	0.007	0.745	1124
Average Maturity	-0.024	0.041	0.719	733
Real Exchange Rate Undervaluation	-0.003	0.005	0.689	2877
Interest Expenditure (% of total Expenditure)	0.007	0.016	0.660	1908
Interest-Growth Differential	-0.004	0.011	0.646	2154
FDI (% of GDP)	0.014	0.038	0.645	2972
Debt owed to Commercial Banks (% of total)	-0.008	0.021	0.638	1423
US TBill Rate	-0.021	0.062	0.630	3592
Nominal GDP Growth	0.002	0.008	0.600	3046
Debt Service due (% of GDP)	-0.006	0.024	0.599	1602
GDP per Capita	0.008	0.043	0.576	3291
Amortisation of Total Public Debt (% of GDP)	-0.002	0.011	0.570	2030
Average Effective Interest Rate	0.000	0.003	0.529	1951
Debt owed to Multilateral Creditors (% of total)	-0.001	0.011	0.526	1632
Short Term Debt (% of FX Reserves)	0.000	0.001	0.518	1984

Note: Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the pooled logit model, with 3 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015.

Table 1.12: Robustness check 4: Random effects logit model

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.302	0.071	1.000	3188
Current Account Balance (% of GDP)	-0.054	0.016	1.000	3275
FX Reserves Growth	-0.010	0.003	0.998	3051
FX Reserves (% of GDP)	-0.035	0.013	0.997	3103
Real GDP Growth	-0.041	0.024	0.960	3269
Primary Balance Gap (% of GDP)	-0.029	0.018	0.950	688
Openness: (M+X)/ GDP	-0.005	0.003	0.948	3254
Overall Fiscal Balance (% of GDP)	-0.055	0.035	0.941	2035
Gross Financing Need	0.019	0.015	0.891	735
Primary Balance (% of GDP)	-0.048	0.041	0.879	1874
Change in Net Claims on Central Government	0.000	0.000	0.861	2415
Unemployment Rate	-0.025	0.024	0.851	2546
Foreign Exchange Debt (% of GDP)	0.007	0.007	0.833	1124
Short Term Debt (% of total)	-0.007	0.008	0.823	2107
LIBOR	0.045	0.050	0.814	3670
FX Reserve Coverage	-0.012	0.015	0.779	1581
Concessional Debt (% of total)	-0.005	0.006	0.775	1632
CPI Inflation	0.004	0.006	0.771	2804
Real Exchange Rate Undervaluation	-0.003	0.004	0.766	2877
Domestic Credit Gap	0.001	0.002	0.757	2588
Real Interest Rate	0.010	0.016	0.735	1387
Short Term External Debt (% of GDP)	-0.005	0.008	0.721	2142
Debt (% of GDP)	-0.002	0.004	0.702	3073
External Debt (% of GDP)	-0.001	0.003	0.682	2475
Average Maturity	-0.017	0.043	0.657	733
FDI (% of GDP)	0.012	0.035	0.636	2972
Interest Expenditure (% of total Expenditure)	0.005	0.017	0.617	1908
Debt owed to Commercial Banks (% of total)	-0.005	0.016	0.614	1423
US TBill Rate	0.012	0.043	0.612	3592
Amortisation of Total Public Debt (% of GDP)	-0.002	0.009	0.595	2030
Short Term Debt (% of FX Reserves)	0.000	0.001	0.585	1984
Average Effective Interest Rate	0.000	0.003	0.554	1951
Debt Service due (% of GDP)	-0.003	0.021	0.551	1602
GDP per Capita	-0.003	0.040	0.526	3291
Nominal GDP Growth	0.000	0.006	0.525	3046
Debt owed to Multilateral Creditors (% of total)	0.000	0.010	0.504	1632
Interest-Growth Differential	0.000	0.008	0.503	2154

Note: Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the random effects logit model, with 2 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015.

Table 1.13: Robustness check 5: Pooled probit model

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.138	0.034	1.000	3188
Current Account Balance (% of GDP)	-0.025	0.007	1.000	3275
FX Reserves (% of GDP)	-0.016	0.005	0.998	3103
FX Reserves Growth	-0.003	0.001	0.995	3051
Primary Balance Gap (% of GDP)	-0.018	0.009	0.973	688
Openness: (M+X)/ GDP	-0.003	0.001	0.969	3254
Real GDP Growth	-0.020	0.011	0.959	3269
Overall Fiscal Balance (% of GDP)	-0.026	0.016	0.954	2035
Foreign Exchange Debt (% of GDP)	0.004	0.003	0.909	1124
Primary Balance (% of GDP)	-0.023	0.018	0.905	1874
Gross Financing Need	0.008	0.007	0.872	735
Short Term Debt (% of total)	-0.004	0.004	0.861	2107
Change in Net Claims on Central Government	0.000	0.000	0.844	2415
Unemployment Rate	-0.010	0.010	0.832	2546
LIBOR	0.020	0.024	0.798	3670
CPI Inflation	0.002	0.003	0.798	2804
FX Reserve Coverage	-0.007	0.008	0.796	1581
Domestic Credit Gap	0.001	0.001	0.781	2588
Concessional Debt (% of total)	-0.002	0.003	0.764	1632
Interest Expenditure (% of total Expenditure)	0.005	0.007	0.764	1908
Short Term External Debt (% of GDP)	-0.003	0.004	0.753	2142
Real Interest Rate	0.005	0.008	0.742	1387
Real Exchange Rate Undervaluation	-0.001	0.002	0.741	2877
Debt (% of GDP)	-0.001	0.002	0.735	3073
External Debt (% of GDP)	-0.001	0.001	0.717	2475
Average Maturity	-0.010	0.020	0.700	733
FDI (% of GDP)	0.005	0.016	0.633	2972
Amortisation of Total Public Debt (% of GDP)	-0.001	0.004	0.626	2030
Debt owed to Commercial Banks (% of total)	-0.002	0.008	0.618	1423
US TBill Rate	0.004	0.020	0.585	3592
Nominal GDP Growth	0.001	0.003	0.581	3046
Short Term Debt (% of FX Reserves)	0.000	0.000	0.576	1984
GDP per Capita	-0.003	0.017	0.570	3291
Debt Service due (% of GDP)	-0.001	0.010	0.555	1602
Debt owed to Multilateral Creditors (% of total)	0.001	0.005	0.544	1632
Interest-Growth Differential	0.000	0.004	0.542	2154
Average Effective Interest Rate	0.000	0.001	0.520	1951

Note: Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the pooled probit model, with 2 additional controls per specification. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015.

Table 1.14: Robustness check 6: Using new measure of output gap

	Coeff.	**SE**	**C(0)**	**Obs.**
Output Gap	0.074	0.018	1.000	3188
Current Account Balance (% of GDP)	-0.025	0.007	1.000	3275
FX Reserves (% of GDP)	-0.016	0.005	0.998	3103
FX Reserves Growth	-0.003	0.001	0.995	3051
Primary Balance Gap (% of GDP)	-0.018	0.009	0.971	688
Openness: (M+X)/ GDP	-0.003	0.001	0.967	3254
Real GDP Growth	-0.019	0.011	0.954	3269
Overall Fiscal Balance (% of GDP)	-0.026	0.016	0.953	2035
Foreign Exchange Debt (% of GDP)	0.004	0.003	0.907	1124
Primary Balance (% of GDP)	-0.023	0.018	0.904	1874
Gross Financing Need	0.008	0.007	0.873	735
Short Term Debt (% of total)	-0.004	0.004	0.859	2107
Change in Net Claims on Central Government	0.000	0.000	0.843	2415
Unemployment Rate	-0.010	0.010	0.835	2546
LIBOR	0.020	0.024	0.800	3670
FX Reserve Coverage	-0.007	0.008	0.797	1581
CPI Inflation	0.002	0.003	0.797	2804
Domestic Credit Gap	0.001	0.001	0.781	2588
Interest Expenditure (% of total Expenditure)	0.005	0.007	0.767	1908
Concessional Debt (% of total)	-0.002	0.003	0.766	1632
Short Term External Debt (% of GDP)	-0.003	0.004	0.753	2142
Real Exchange Rate Undervaluation	-0.001	0.002	0.744	2877
Real Interest Rate	0.005	0.008	0.738	1387
Debt (% of GDP)	-0.001	0.002	0.732	3073
External Debt (% of GDP)	-0.001	0.001	0.716	2475
Average Maturity	-0.010	0.020	0.695	733
FDI (% of GDP)	0.006	0.016	0.639	2972
Amortisation of Total Public Debt (% of GDP)	-0.001	0.004	0.628	2030
Debt owed to Commercial Banks (% of total)	-0.002	0.008	0.613	1423
US TBill Rate	0.005	0.020	0.589	3592
Nominal GDP Growth	0.001	0.003	0.581	3046
Short Term Debt (% of FX Reserves)	0.000	0.000	0.575	1984
GDP per Capita	-0.003	0.017	0.569	3291
Debt Service due (% of GDP)	-0.001	0.010	0.556	1602
Debt owed to Multilateral Creditors (% of total)	0.001	0.005	0.543	1632
Interest-Growth Differential	0.000	0.004	0.539	2154
Average Effective Interest Rate	0.000	0.001	0.522	1951

Reported are estimation results from the EBA regression. The dependent variable is Baldacci et al. (2011) definition of fiscal distress: (i) debt default or restructuring, (ii) sovereign bond yield pressure, (iii) large IMF-supported program, and (iv) excessive inflation. Estimations are performed using the pooled Logit model, with 2 additional controls per specification. The output gap is generated using a one-sided Christiano-Fitzgerald filter. The variables are ranked according to their robustness, with most robust indicators placed on top. The sample covers 29 advanced and 52 emerging economies for the period 1970-2015.

CHAPTER 2

Bayesian Structural VAR models: a new approach for prior beliefs on impulse responses

2.1 Introduction

Structural Vector Autoregressive models (SVARs) are extensively used in applied Macroeconomics. To provide results that can be interpreted economically, SVARs require identifying restrictions. It has become popular to introduce identifying restrictions in the form of sign restrictions on selected structural parameters. This is typically done using a Bayesian approach with informative prior beliefs that reflect the intended signs (Uhlig, 2005, Baumeister and Hamilton, 2015, Arias et al., 2018).

Implementing sign restrictions presents the researcher with a trade-off. There exist infinitely many prior probability distributions that reflect a desired set of sign restrictions. Out of this large class of priors, the literature often limits the analysis to the Normal-inverse-Wishart-Uniform prior (hereafter NiWU) in order to simplify the analysis of the posterior distribution (Uhlig, 2005, Rubio-Ramirez et al., 2010). However, this constrains the type of prior information introduced by the researcher to the one that can be modelled by the Normal-inverse-Wishart-Uniform prior. This is an important limitation, given that, even in a large sample, the results are affected by the specific probability distribution used to model the desired sign restrictions. Yet, moving beyond the NiWU prior makes the posterior distribution (and hence the results) more challenging to analyse (Arias et al., 2018). A trade-off hence emerges between the flexibility in the selection of the prior distribution used, advocated by Baumeister and Hamilton (2015), and the tractability of the posterior distribution, favoured by Rubio-Ramirez et al. (2010).

The first contribution of the paper consists in developing a methodology that makes the above trade-off disappear. We build our methodology on a new importance sampler

This chapter is based on joint work with Michele Piffer.

that uses the posterior distribution of the convenient NiWU case as an importance distribution. While relatively unchallenging to implement, importance samplers require that the importance distribution covers the relevant support of the target distribution (Creal, 2012). When working directly on structural parameters, this condition can be argued to hold only for prior beliefs that do not differ considerably from the NiWU prior, a case explored by Arias et al. (2018). We build on their work and show that this condition holds for a much wider class of prior beliefs if one builds the importance sampler in two separate stages: first on the reduced form parameters, and second on the mapping into structural parameters. We show that, after acknowledging this point, the trade-off mentioned above disappears: one can follow Baumeister and Hamilton (2015) and use prior beliefs that differ considerably from the one implied in the NiWU approach, but the sampling of the corresponding posterior distribution does not become technically more involved compared to the techniques developed by Rubio-Ramirez et al. (2010) for the NiWU approach. Accordingly, the methodology offers the most desirable scenario, as it allows for prior flexibility at no additional computational cost. To further confirm the effectiveness of our sampler, we show that the results of the applications in this paper are the same when exploring the posterior distribution using the sequential approach by Waggoner et al. (2016), which is more time-demanding but also suitable to explore potentially ill-shaped distributions. We first develop our methodology by focusing on the case of only sign restrictions, and then discuss an extension that combines sign and zero restrictions.

The second contribution of the paper consists in proposing a new approach for sign restrictions on impulse responses, which are arguably the most important statistic of SVAR models. On the one hand, starting from prior beliefs directly on impulse responses makes it technically demanding to explore the posterior distribution (see Kociecki, 2010, Barnichon and Matthes, 2018 and Plagborg-Møller, 2019). On the other hand, as discussed above, the use of the NiWU approach reduces the flexibility on the actual prior probability distribution introduced on the parameters of interest. We propose a compromise that parametrizes the structural VAR model as in Uhlig (2005), hence in the reduced form autoregressive elements and in the *contemporaneous* impulse responses. We then depart from Uhlig (2005) and do not restrict the prior on the contemporaneous impulse responses to the one implied by the NiWU prior. Instead, we allow for a general prior distribution. In offering prior flexibility on the impulse response horizon where flexibility is needed the most (Canova and Pina, 2005 and Canova and Paustian, 2011), our approach provides a balance between prior flexibility on the key structural parameters, and conditionally conjugate priors on all the

remaining parameters. We then explore the posterior distribution of the remaining parameters using the newly developed importance sampler.

Having developed a tractable framework that can handle a wide class of prior distributions on the contemporaneous impulse responses, we illustrate that indeed the results in applied work can be sensitive to the prior distribution used. When mapping reduced form parameters into structural parameters, the criterion used in the NiWU approach focuses on orthogonal matrices, namely that orthogonal matrices are conditionally uniformly distributed. This approach can unintentionally treat as equally plausible orthogonal matrices that imply an impact of a one-standard-deviation shock as big as a multiple of the standard deviation of a variable of the model. We propose a prior specification that ensures that the prior mass associated with one-standard-deviation shocks is in line with the scaling of the variables, in a way modelled explicitly by the researcher through a training sample and a set of hyperparameters. We show that this new feature can tighten posterior bands considerably, potentially leading to new results in applied work. Compared to the NiWU approach, the tighter posterior bands do not trivially come from tighter priors (and indeed we show that the opposite holds). They come from the fact that the mapping from reduced form to structural parameters is made consistent with the volatility of the variables. Alternative prior specifications are, of course, possible. All in all, the paper suggests that prior beliefs on structural parameters should be selected carefully, as advocated by Baumeister and Hamilton (2015), but also that the NiWU approach advocated by Rubio-Ramirez et al. (2010) offers the required point of departure to explore the posterior distribution associated with this more general approach.

Since the traditional NiWU approach to sign restricted SVARs frequently implies relatively wide posterior bands on impulse responses, many studies have proposed to combine sign restrictions with additional restrictions on other statistics (see, for example, Kilian and Murphy, 2012, Antolín-Díaz and Rubio-Ramírez, 2018 and Amir-Ahmadi and Drautzburg, 2018). We argue that taking into account the scaling of the variables when forming prior beliefs to model sign restrictions on impulse responses is sufficient to deliver sharper inference, to the point that no additional restriction is needed to interpret the results. We show this by applying our methodology to the long lasting debate on what drives the unexpected variations in the price of oil and the associated effects on the US economy. Kilian and Murphy (2012) address this question using sign restrictions on contemporaneous impulse responses applied in a setting close to the NiWU approach. They show that sign restrictions alone deliver posterior bands that are too wide to disentangle the different channels driving oil price dynamics. They propose to add restrictions on the elasticity of oil supply, and find

that oil demand shocks are the main drivers of oil price dynamics. We show, instead, that applying the same initial sign restrictions in a way that is more consistent with the scaling of the variables can tighten posterior bands considerably without need for restrictions on elasticities (as in Kilian and Murphy, 2012) nor on estimated shocks and historical decompositions (as in the extension of the model by Kilian and Murphy, 2012 proposed by Antolín-Díaz and Rubio-Ramírez, 2018).

More precisely, we construct our application to the oil market as follows. We use a prior probability distribution that treats different structural shocks symmetrically, ensuring that the prior does not favour one shock over the other as drivers of the variables in the model. We then show that the wide posterior bands in Kilian and Murphy (2012) can be replicated using prior beliefs that attach 80% prior probability mass to very strong effects of one-standard-deviation shocks on the variables of the model, based on an initial quantitative assessment from a training sample. Last, we tighten the prior mass by making it more consistent with the scaling of each variable of the model. While confirming the initial results by Kilian and Murphy (2012) on the importance of oil demand shocks, we find that oil supply shocks have a considerable effect on oil price dynamics. Quantitatively, we find that as much as 30-40% of the forecast error variance of the real price of oil can be explained by oil supply shocks. Our results confirm the findings by Baumeister and Hamilton (2019) and Caldara et al. (2018). We also find that oil supply shocks were indeed the prevailing driver of the drop in oil production during the first Gulf War, a feature that Antolín-Díaz and Rubio-Ramírez (2018) introduce as an identifying restriction.

From the methodological point of view, we complement the work by Sims and Zha (1998) and Baumeister and Hamilton (2015) and study the case of beliefs on contemporaneous impulse responses rather than on the contemporaneous relation among variables. Baumeister and Hamilton (2018) combine prior beliefs on contemporaneous relations and contemporaneous impulse responses. Relative to Baumeister and Hamilton (2018), we focus on impulse responses and propose a different prior specification and posterior sampler. Last, we relate to Giacomini and Kitagawa (2015) in stressing the mapping from reduced form to structural parameters, but we concentrate on a single prior.

The paper is organized as follows. Section 2.2 outlines the methodology proposed and discusses its relation to the existing literature. Section 2.3 shows an illustrative example on simulated data based on the estimated bivariate VAR model by Baumeister and Hamilton (2015). Section 2.4 reports the application to the oil market. Section 2.5 concludes.

2.2 The methodology

In this section we present the structural VAR model and summarize the traditional NiWU approach to sign restrictions. We then outline our methodology and discuss the new importance sampler. Last, we propose one possible prior distribution that can be used with our approach. Our importance sampler can be used also with other prior beliefs.

2.2.1 The model

Following Uhlig (2005), we write the structural VAR model as

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\pi}_0 + \sum_{l=1}^{p} \Pi_l \boldsymbol{y}_{t-l} + B\boldsymbol{\epsilon}_t, \\ &= \Pi \boldsymbol{w}_t + B\boldsymbol{\epsilon}_t, \qquad\qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I_k), \end{aligned} \tag{2.1}$$

where $\boldsymbol{y}_t$ is a $k \times 1$ vector of endogenous variables, $\boldsymbol{\epsilon}_t$ is a $k \times 1$ vector of structural shocks, and $\boldsymbol{w}_t = (1, \boldsymbol{y}'_{t-1}, .., \boldsymbol{y}'_{t-p})'$ is an $m \times 1$ vector of the constant and p lags of the variables, with $m = kp + 1$. The matrix $\Pi = [\boldsymbol{\pi}_0, \Pi_1, .., \Pi_p]$ is of dimension $k \times m$. We normalize the covariance matrix of $\boldsymbol{\epsilon}_t$ to the identity matrix.[1]

Matrix B in equation (2.1) captures the contemporaneous effects of one-standard-deviation shocks, while future horizons of the impulse responses are calculated using model (2.1) recursively. Although structural VARs can also be specified in matrix $A = B^{-1}$ rather than in B (see, for example, Sims and Zha, 1998), we use model (2.1) as in Uhlig (2005) in order to emphasize the key objects of interest for our analysis, which are the contemporaneous impulse responses. We focus on the case in which the researcher expresses identifying restrictions in the form of sign (and possibly zero) restrictions on contemporaneous impulse responses.[2]

[1] This normalization is frequently used in applications that employ sign restrictions on impulse responses, see for example Canova and De Nicoló (2002), Uhlig (2005), Benati and Surico (2009).

[2] Whether the model is more conveniently expressed in $A = B^{-1}$ or B (or even in a combined form) depends on whether the identifying restrictions introduced by the researcher are more naturally expressed on contemporaneous relation among variables or contemporaneous effects of the shocks, respectively. For example, the literature on the identification of monetary policy shocks employs restrictions either on B, as in Uhlig (2005), or on A, as in Arias et al. (2019) and Baumeister and Hamilton (2018). Restrictions imposed on one form might not be apparent in the other form, due to the nonlinearities in the mapping from one to another. Going through the publications of all top-five journals and the Journal of Monetary Economics since 1998, we found that around 13% of the total number of issues checked included at least one application of Structural Vector Autoregressive models. Of the total number of SVAR applications that we found, approximately 15% specifies the model in the A form, 76% specifies the model in the B form, and 9% specifies the model in the hybrid AB form. The detailed list is available at this link.

The reduced form representation of the structural model is

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\pi}_0 + \sum_{l=1}^{p} \Pi_l \boldsymbol{y}_{t-l} + \boldsymbol{u}_t, \\ &= \Pi \boldsymbol{w}_t + \boldsymbol{u}_t, \qquad\qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma), \end{aligned} \tag{2.2}$$

where it holds that $\boldsymbol{u}_t = B\boldsymbol{\epsilon}_t$ and $\Sigma = BB'$. Orthogonal matrices Q, which by definition satisfy $QQ' = I_k$, allow for the mapping from reduced form to structural parameters, with

$$B = h(\Sigma)Q, \tag{2.3}$$

and $h(\Sigma)$ a factorization of Σ satisfying $h(\Sigma)h(\Sigma)' = \Sigma$, for example the Cholesky factorization.

2.2.2 The NiWU approach used in the literature

The most popular approach for sign restricted SVAR models expresses prior beliefs on the parameters $(\boldsymbol{\pi}, \Sigma, Q)$, with $\boldsymbol{\pi} = vec(\Pi)$ the $km \times 1$ vector that stacks the columns of Π. As already discussed in the literature, when $p(\boldsymbol{\pi}, \Sigma)$ falls within either the independent or the conjugate Normal-inverse-Wishart prior, drawing from the joint posterior distribution $p(\boldsymbol{\pi}, \Sigma|Y)$ is technically convenient (see, for example, Koop et al., 2010). One can then extract Q matrices uniformly in the parameter space $\mathcal{Q}_\Sigma$, defined as the set of orthogonal matrices such that the sign restrictions on the structural parameters are satisfied, given a draw of Σ. Draws of Q are retained if the sign restrictions are satisfied, and are discarded otherwise.

The convenience of the NiWU approach is that efficient algorithms exist for the sampling of the posterior distribution, developed for example by Rubio-Ramirez et al. (2010). In addition, the possibility of discarding undesired draws allows for the straightforward introduction of sign restrictions not only on contemporaneous impulse responses, but also on future horizons. The inconvenience is that the prior probability distribution is not directly specified on the structural parameters of interest, namely the impulse responses, but on reduced form parameters and on orthogonal matrices. Since impulse responses are not point identified parameters (or in the terminology by Rubio-Ramirez et al., 2010, are not exactly identified parameters), the implicit prior distribution matters also in a large sample and must be selected carefully (Baumeister and Hamilton, 2015).

To appreciate the importance of the above point, consider for simplicity the case of sign restrictions on the contemporaneous impulse responses. The restrictions can be modelled with a probability distribution $p(B)$ that attaches zero mass to the values that

do not satisfy the restrictions. However, there are infinitely many probability distributions $\{p(B)_1, p(B)_2, p(B)_3, ...\}$ that reflect the same candidate sign restrictions. Since B is not point identified, the posterior distributions $\{p(B|Y)_1, p(B|Y)_2, p(B|Y)_3, ...\}$ differ even in a large sample. Accordingly, not only the sign restrictions are important, but also the actual probability distribution used to model them (Baumeister and Hamilton, 2015). Under the NiWU approach the flexibility on $p(B)$ is constrained by the fact that it is expressed indirectly through $p(\Sigma, Q)$, that $p(\Sigma)$ must be the inverse Wishart probability distribution, and that $p(Q|\Sigma)$ is uniform in the space $\mathcal{Q}_\Sigma$.[3]

2.2.3 The normal p(B) approach proposed in this paper

To overcome the limitations discussed in the previous section, we propose to express prior beliefs directly on B. We then develop an importance sampler that ensures that the additional flexibility on the prior specification does not come at a computational cost.

2.2.3.1 Prior beliefs expressed directly on (π, B)

We parametrize the model as in equation (2.1) and express prior beliefs on $(\boldsymbol{\pi}, B)$, i.e.

$$p(\boldsymbol{\pi}, B) = p(\boldsymbol{\pi}|B) \cdot p(B). \tag{2.4}$$

Since $\boldsymbol{\pi}$ is identified, $p(\boldsymbol{\pi})$ matters less compared to $p(B)$, as long as the sample is sufficiently long. Hence, as also in the NiWU approach, we restrict $p(\boldsymbol{\pi})$ to

$$\boldsymbol{\pi} \sim N(\boldsymbol{\mu}_\pi, V_\pi), \tag{2.5}$$

where $\boldsymbol{\mu}_\pi$ and V_π can be a function of B. By contrast, $p(B)$ allows for a large class of prior distributions, granting the researcher flexibility on the prior beliefs used to express sign restrictions on key structural parameters.[4]

[3] That prior beliefs on one parametrization imply questionable or unintended features on some other parametrization is to some extent inevitable. Baumeister and Hamilton (2015) argue that prior beliefs should be judged relative to the structural parametrization of interest, which in our application is B. Arias et al. (2018) derive analytically the distribution implied by the NiWU approach on structural parameters. See also Section 2.A.2 in the Appendix.

[4] As in Baumeister and Hamilton (2015) and Baumeister and Hamilton (2019), we require that $p(B)$ is everywhere nonnegative, and when integrated over the set of all values of B, it produces a finite positive number. If the posterior distribution is then explored with our importance sampler, an additional requirement is that the variance of the weights in Stage A of our algorithm is finite (Geweke, 1989). We return to this point below as well as in Section 2.A.3.2 of the Appendix.

As we show in Section 2.A.3 of the Appendix, the joint posterior distribution satisfies

$$p(\boldsymbol{\pi}, B|Y) = p(\boldsymbol{\pi}|B, Y) \cdot p(B|Y), \tag{2.6}$$

where

$$\boldsymbol{\pi}|B, Y \sim N(\boldsymbol{\mu}_\pi^*, V_\pi^*), \tag{2.7}$$

$$p(B|Y) \propto p(B) \cdot |det(B)|^{-T} \cdot |det(V_\pi)|^{-\frac{1}{2}} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot \\ \cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T \otimes (BB')^{-1}\right)\tilde{\boldsymbol{y}} - \boldsymbol{\mu}_\pi^{*\prime} V_\pi^{*-1} \boldsymbol{\mu}_\pi^* + \boldsymbol{\mu}_\pi' V_\pi^{-1} \boldsymbol{\mu}_\pi\right\}}, \tag{2.8}$$

with $\tilde{\boldsymbol{y}}$, W, $\boldsymbol{\mu}_\pi^*$ and V_π^* defined in the Appendix. Drawing from $p(\boldsymbol{\pi}, B|Y)$ requires a suitable posterior sampling procedure for the k^2 elements in $p(B|Y)$, or even for fewer parameters in case zero restrictions are introduced on B. Draws for the km elements in $\boldsymbol{\pi}|B, Y$ can instead be obtained with a standard random number generator.

The above approach strikes a balance between flexibility and tractability. On the one hand, it grants the researcher flexibility on impulse responses at the horizon where flexibility is needed the most, which is the horizon of the impact effect. On the other hand, as also the NiWU approach, it makes the analysis more tractable by using a normal prior distribution on $\boldsymbol{\pi}$. The normality on $\boldsymbol{\pi}$ is not restrictive except in small samples, given that $\boldsymbol{\pi}$ is point identified. Since sign restrictions on impulse responses are frequently introduced contemporaneously rather than on future horizons, we do not view our framework as particularly restrictive. In addition, by parametrizing the model in $\boldsymbol{\pi}$, our approach makes it straightforward to use the prior by Litterman (1986) (which is applied directly on $\boldsymbol{\pi}$), simplifying the analysis compared to Sims and Zha (1998) and Baumeister and Hamilton (2015).

2.2.3.2 A new posterior sampler for $p(B|Y)$

To make our approach viable in applied work we require an efficient algorithm that explores the posterior distribution $p(B|Y)$ from equation (2.8). When prior beliefs $p(B)$ take the special case implied by the NiWU approach, the posterior distribution $p(B|Y)$ can be explored using existing algorithms for the NiWU approach (Section 2.A.2 of the Appendix). We now develop an extension of such algorithms to allow for a wider class of prior beliefs on B.

We build our sampling procedure on importance sampling techniques. Consider a parameter vector of interest, $\boldsymbol{\theta}$. Suppose we are interested in sampling from the target distribution $p(\boldsymbol{\theta})^{target}$, and suppose we cannot draw from $p(\boldsymbol{\theta})^{target}$ directly, but can evaluate it. In addition, suppose that we can extract proposal draws from the

importance function $p(\boldsymbol{\theta})^{importance}$. To the extent that the importance function fully covers the support of $p(\boldsymbol{\theta})^{target}$, we can obtain draws from $p(\boldsymbol{\theta})^{target}$ by resampling with replacement the draws $\{\boldsymbol{\theta}_i\}$ obtained from the importance distribution using weights $w(\boldsymbol{\theta}_i) = \frac{p(\boldsymbol{\theta}=\boldsymbol{\theta}_i)^{target}}{p(\boldsymbol{\theta}=\boldsymbol{\theta}_i)^{importance}}$ (see for example Koop, 2003, chapter 4). A popular diagnostic metric is the effective sample size $ESS = \left(\sum_i \left(w_i / \sum_i (w_i)\right)^2\right)^{-1}$, which captures the effective number of draws used to represent the target probability, given an initial number of proposal draws. If the importance function sufficiently covers the support of the target function, a small effective sample size suggests increasing the number of draws from the importance function. If, instead, we cannot ensure that the importance function gives sufficient mass to the support of the target function, the importance function must be changed irrespectively of the effective sample size (see the simulation exercise in Section 2.3).

Define $p(B|Y)_{Np(B)}$ as the posterior distribution associated with the general prior $p(B)$ from our approach, which we denote Np(B) (equation 2.8), and $p(B|Y)_{NiWU}$ as the posterior distribution associated with the NiWU approach. Since sampling from $p(B|Y)_{NiWU}$ is not challenging, in principle one could set $\boldsymbol{\theta} = B$, $p(\boldsymbol{\theta})^{target} = p(B|Y)_{Np(B)}$ and $p(\boldsymbol{\theta})^{importance} = p(B|Y)_{NiWU}$. Arias et al. (2018) show that this approach works successfully if the target distribution $p(B|Y)_{Np(B)}$ does not differ too much from the tractable distribution $p(B|Y)_{NiWU}$. However, this procedure does not work in a general framework, because one cannot ensure that $p(B|Y)_{NiWU}$ sufficiently covers the support of $p(B|Y)_{Np(B)}$, except for special cases in which $p(B)_{Np(B)}$ is close to the prior on B implied in the NiWU approach.

We circumvent the above challenge by exploring $p(B|Y)_{Np(B)}$ indirectly. First, define the following functions:

- $p(\Sigma|Y)_{Np(B)}$: posterior distribution on Σ implied by $p(B|Y)_{Np(B)}$ from equation (2.8), corresponding to our Np(B) approach which expresses the prior on structural parameters;
- $p(\Sigma|Y)_{NiWU}$: posterior distribution on Σ corresponding to the NiWU approach which expresses priors on reduced form parameters and rotation matrices;
- $p(Q|\Sigma)_{Np(B)}$: conditional distribution on Q implicit in the prior $p(B)$ from our Np(B) approach;
- $p(Q|\Sigma)_{NiWU}$: conditional distribution on Q employed in the NiWU approach, which coincides with a uniform distribution on the space $\mathcal{Q}_\Sigma$;
- $p(B)_{Np(B)}$: prior distribution on B used in our Np(B) approach.

Then, notice that drawing from $p(B|Y)_{Np(B)}$ is equivalent to drawing from $p(\Sigma|Y)_{Np(B)}$ and mapping Σ into B using draws of Q from $p(Q|\Sigma)_{Np(B)}$. Accordingly, consider the following importance sampling procedure. First, explore $p(\Sigma|Y)_{Np(B)}$ using $p(\Sigma|Y)_{NiWU}$ as an importance function. Since Σ is point identified, $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$ are close to each other except in small samples, making $p(\Sigma|Y)_{NiWU}$ a candidate importance function for $p(\Sigma|Y)_{Np(B)}$. Then, use $p(Q|\Sigma)_{NiWU}$ as a proposal function for $p(Q|\Sigma)_{Np(B)}$ to map draws from $p(\Sigma|Y)_{Np(B)}$ into draws from $p(B|Y)_{Np(B)}$. Since $p(Q|\Sigma)_{NiWU}$ is conditionally uniform, it fully explores the parameter space $\mathcal{Q}_\Sigma$, reducing to zero the probability that $p(Q|\Sigma)_{NiWU}$ does not explore the relevant parameter space covered by $p(Q|\Sigma)_{Np(B)}$. In the first stage, a low effective sample size suggests that the sample is too short to imply that $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$ are similar distributions, a conjecture that can be verified indirectly by computing the effective sample size and employing existing diagnostic procedures. In the second stage, a low effective sample size only suggests to increase the number of draws from the importance function.

Section 2.A.3.2 of the Appendix provides a further discussion of the sampler. It gives the analytical form for $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$, and shows that numerically evaluating $p(Q|\Sigma)_{Np(B)}$ only requires evaluating the prior $p(B)_{Np(B)}$. When only sign restrictions are introduced on B, when $\boldsymbol{\mu}_\pi$ and V_π are not a function of B, and when the NiWU employed to obtain proposal draws is used in its independent prior specification, the sampler can then be implemented using the following algorithm:

Our Algorithm (sign restrictions):

Stage A: generate draws from $p(\Sigma|Y)_{Np(B)}$:

1. run a Gibbs sampler to explore $p(\boldsymbol{\pi}, \Sigma|Y)_{NiWU}$ using m_1 burn-in replications and m_2 retained replications. Store the retained draws in $\{\Sigma_d\}_{d=1}^{m_2}$, which by construction represent draws from $p(\Sigma|Y)_{NiWU}$;
2. for each Σ_d compute weights

$$w_d^{\text{stage A}} = \frac{p(\Sigma = \Sigma_d|Y)_{Np(B)}}{p(\Sigma = \Sigma_d|Y)_{NiWU}} \propto \frac{\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB}{|det(\Sigma_d)|^{-\frac{d+k}{2}} \cdot e^{-\frac{1}{2}tr[\Sigma_d^{-1}S]}}, \qquad (2.9)$$

with $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB$ the integral of $p(B)_{Np(B)}$ along the space of B that implies $BB' = \Sigma_d$. Assess if the effective sample size $ESS^{\text{A}} = \left(\sum_d \left(w_d^{\text{stage A}} / \sum_d (w_d^{\text{stage A}})\right)^2\right)^{-1}$ is sufficiently high[5];

[5] An ESS below 0.5 suggests that the weights are unsuitable for re-sampling. See Appendix 2.A.3.2 for a discussion.

3. randomly select Σ_d from $\left\{\Sigma_d\right\}_{d=1}^{m_2}$ with replacement using weights $w_d^{\text{stage A}}$ to generate draws from $p(\Sigma|Y)_{Np(B)}$;

Stage B: map draws from $p(\Sigma|Y)_{Np(B)}$ into draws from $p(B|Y)_{Np(B)}$:

4. draw an orthogonal matrix (Q_d) using the method by Rubio-Ramirez et al. (2010) and compute $B_d = h(\Sigma_d)Q_d$;

5a. if (B_d) satisfies the sign restrictions, store (B_d, Q_d) and proceed to Step 6;

5b. if (B_d) does not satisfy the sign restrictions, repeat Step 4 up to m_4 times. Stop as soon as (B_d) satisfies the sign restrictions and proceed to Step 6, otherwise discard (Σ_d) and move back to Step 3;

6. repeat Steps 3 to 5 until m_5 draws are stored;

7. for all draws $\left\{B_d, Q_d\right\}_{d=1}^{m_5}$ compute weights

$$w_i^{\text{stage B}} = \frac{p(Q = Q_d|\Sigma_d)_{Np(B)}}{p(Q = Q_d|\Sigma_d)_{NiWU}} \propto p(B = B_d). \tag{2.10}$$

Assess if the effective sample size $ESS^{\text{B}} = \left(\sum_i \left(w_d^{\text{stage B}} / \sum_i (w_i^{\text{stage B}})\right)^2\right)^{-1}$ is larger than a desired minimum number m_6. If so, proceed to Step 8, otherwise move back to Steps 3 to 6 and increase m_5;

8. generate $\left\{B_d\right\}_{d=1}^{ESS^{\text{B}}}$ by resampling the draws $\left\{B_d\right\}_{d=1}^{m_5}$ from Step 7 with replacement using weights $w_d^{\text{stage B}}$.[6]

Our algorithm resamples the posterior draws from the NiWU approach and makes them representative of the posterior distribution associated with the generic prior beliefs $p(B)$ from our approach. In the rest of the paper we document that the sampling time of our algorithm is roughly 30 minutes in the oil application and about 6 minutes in the simulation exercise, depending on the sample size. In Section 2.A.3.2 of the Appendix we also argue that the size of the dataset beyond which the effective sample size in Stage A is sufficiently high is relatively small, further suggesting that our algorithm can be used in samples frequently used in applied work. The only computationally demanding term to evaluate for our algorithm is $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB$, which we evaluate numerically as discussed in Section 2.A.3 of the Appendix. Yet, the Appendix shows that an approximate algorithm that sets $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB = 1$ reaches an almost identical approximation of the posterior distribution, further reducing the computational time.

[6]Since ESS^{B} can be much smaller than m_5, we resample $\left\{B_d\right\}_{d=1}^{m_5}$ only ESS^{B} times (or the closest integer), rather than m_5 times. This avoilds unnecessary repetitions.

Our algorithm offers a way of implementing sign restrictions. Only two modifications are required to extend the algorithm to account for also zero restrictions on B. First, the computation of $p(\Sigma|Y)_{Np(B)}$ (required for the weights in Stage A, Step 2) and the evaluation of $p(Q|\Sigma)_{Np(B)}$ (required for the weights in Stage B, Step 7) must now account for the fact that the mapping from B to Σ features zero restrictions. Accordingly, a numerical approach must be used to compute the corresponding Jacobian transformation, and can be done, for example, using the method developed by Arias et al. (2018). Second, the algorithm generating candidate Q matrices (required in Stage A for the computation of $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)}dB$ and in Stage B to generate proposal draws for Q), must now be replaced with the methods by either Binning (2013) or Arias et al. (2018). The existing version of the algorithm can then be applied to the case in which zero restrictions are introduced on one structural shock of interest.[7]

To further assess whether the algorithm correctly samples from the posterior, we also explore $p(B|Y)_{Np(B)}$ using the Dynamic Striated Metropolis-Hastings algorithm by Waggoner et al. (2016). This alternative algorithm is computationally more demanding, but can handle potentially irregularly shaped posterior distributions and a large number of parameters. Using the posterior distribution from this algorithm, we use simulations to show that the sampling procedure proposed in this section does a good job in exploring $p(B|Y)_{Np(B)}$ even in relatively small samples. Section 2.A.4 of the Appendix discusses how we implement the algorithm by Waggoner et al. (2016).

2.2.4 Proposing one possible prior $p(B)$

The paper has so far developed an approach that uses a general prior distribution $p(B)_{Np(B)}$ for the contemporaneous impulse responses, while still allowing for fast and efficient posterior sampling. We conclude the section on the methodology by discussing one possible prior specification for $p(B)$. Other prior beliefs are also possible, and must ultimately be chosen by the applied researcher.

Specifying prior beliefs $p(B)_{Np(B)}$ is challenging, because the literature still provides limited guidance on explicit prior beliefs on structural parameters. Baumeister and Hamilton (2015) impose restrictions on B^{-1} rather than on B and use the existing literature to form prior beliefs on the contemporaneous elasticities among variables.

[7] In our algorithm, extracting Q matrices from the algorithm by Rubio-Ramirez et al. (2010) ensures that the orthogonal parameter space $\mathcal{Q}$ is explored uniformly, and hence is fully explored. When zero restrictions are introduced, the uniformity in the extraction of Q is lost, except in the case in which zero restrictions are introduced on only one structural shock of interest (Arias et al., 2018). In this case, the full relevant orthogonal parameter space is still explored. When zero restrictions are introduced on the effects of more than one shock of interest, the distribution $p(Q|\Sigma)_{NiWU}$ must be evaluated numerically, and the possibility that the relevant part of the orthogonal space is not explored must be addressed.

However, as discussed by Kilian and Lütkepohl (2017), researchers may lack explicit prior information on the contemporaneous relationship among variables. Instead, they frequently have prior beliefs that do not go beyond the sign of contemporaneous impulse responses. As an example, one may entertain the belief that an exogenous, one-standard-deviation monetary increase in the interest rate decreases inflation, but lacks prior beliefs on the scale of such a decrease.

To overcome this challenge, we propose a prior specification for $p(B)_{Np(B)}$ that builds on a conventional prior specification used in the literature for $p(\boldsymbol{\pi})$ known as the Minnesota prior (see, for example, the discussion in Canova, 2007 and Kilian and Lütkepohl, 2017). The crucial step is to take a stand on what is considered a reasonable scale, or magnitude, for the parameters. With the Minnesota prior, one first associates each variable with a reasonable scale capturing the volatility of the variables. This is usually implemented by estimating the variance σ_i of the residual on univariate AR processes on each variable, using a training sample. Then, Bayesian shrinkage is introduced through a set of hyperparameters that shrink the parameters in $\boldsymbol{\pi}$ towards the random walk or the white noise process, taking the relative scale of the variables into account.

We propose to extend the above procedure as follows. Call b_{ij} the entry of B capturing the effect of a one-standard-deviation shock j to variable i. It can be shown that the covariance restrictions $\Sigma = BB'$ imply

$$-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5}, \tag{2.11}$$

with Σ_{ii} the $i-th$ element of the diagonal of Σ.[8] Accordingly, $\gamma_i = \hat{\Sigma}_{ii}^{0.5}$ provides a candidate assessment of the upper bound for b_{ij}, where $\hat{\Sigma}$ is an estimate based on a training sample. We then introduce two hyperparameters ψ_1 and ψ_2 that control for the location and the spread of $p(b_{ij})$. We use independent normal distributions $N(\mu_{ij}, \sigma_{ij})$ as follows:

1. if no sign restriction is imposed on b_{ij}, set $\mu_{ij} = 0$ and $\sigma_{ij} = \psi_2\gamma_i/1.96$, so that the distribution is symmetric around 0, and 95% of the prior mass is in the space $(-\psi_2\gamma_i,\ \psi_2\gamma_i)$;

2. if b_{ij} is restricted to be positive, start from a normal distribution with $\mu_{ij} = \psi_1\gamma_i$ and calibrate the variance such that the distribution has 95% prior mass in the space $(0,\ \psi_2\gamma_i)$;

[8] Given $\Sigma = BB'$, the equations corresponding to the diagonal elements of Σ are $\Sigma_{ii} = b_{i1}^2 + b_{i2}^2 + ... + b_{ik}^2$. Since Σ_{ii} is nonnegative and since $b_{ij}^2 \geq 0$, each element b_{ij} must satisfy $-\Sigma_{ii}^{0.5} \leq b_{ij} \leq \Sigma_{ii}^{0.5}$. See also equation (33) in Baumeister and Hamilton (2015).

3. if b_{ij} is restricted to be negative, start from a normal distribution with $\mu_{ij} = -\psi_1\gamma_i$ and calibrate the variance such that the distribution has 95% prior mass in the space $(-\psi_2\gamma_i,\ 0)$.

Put differently, since the mode of $p(b_{i,j})$ equals $\psi_1\gamma_i$ (if $b_{i,j}$ is restricted to be positive) or $-\psi_1\gamma_i$ (if $b_{i,j}$ is restricted to be negative), ψ_1 controls for the first moment of the prior. The hyperparameter ψ_2 then controls for the second moment of the prior, given that ψ_2 is positively related to the probability mass attached to $|b_{i,j} > \gamma_i|$. The convenience of the above approach is that the researcher sets a plausible upper bound for the effect of the shocks by selecting γ_i, and then explicitly introduces Bayesian shrinkage through the hyperparameters ψ_1 and ψ_2. ψ_1 and ψ_2 can be treated hierarchically, further providing flexibility on the prior distribution used. If sign restrictions do not identify all shocks, we suggest to numerically introduce the restriction that the non-identified shocks do not replicate the sign restrictions of the identified shocks.[9] Alternative specifications are also possible.

2.3 An illustrative example

In this section we outline the intuition for our approach using simulations on a bivariate VAR model. We then discuss what drives the difference between the Np(B) approach proposed in this paper and the traditional NiWU approach used in the literature.

2.3.1 Simulation exercise

We build the simulation exercise on the model estimated by Baumeister and Hamilton (2015). We first employ ordinary least squares to estimate their bivariate reduced form VAR model, which uses data on the growth rates of the US real labour compensation and total employment from 1970Q1 through 2014Q4, adding a constant and 8 lags. We then use the estimated reduced form VAR as the data generating process. We generate a dataset of 680 draws initializing the data from the estimated unconditional mean. We discard the first 100 draws to make the data less dependent on the initial point, and store the next 100 draws to use as a training sample. We then divide the remaining 480 draws into five pseudo datasets, including up to the first 30, 60, 120, 240 and 480 observations. We use the same training sample for all datasets to improve the comparison, and to avoid an unreasonably short training sample for the dataset of smaller size.

[9] If more than one shock is non-identified, uniqueness could for example be achieved by imposing in addition that the non-identified shocks have distinct sign patterns.

For each pseudo dataset, we estimate the structural VAR model from equation (2.1) by introducing sign restrictions on the contemporaneous impulse responses. We identify the demand shock and the supply shock as the structural shocks that move wages and employment in the same and in the opposite direction, respectively. While the models employ the same sign restrictions, we model such restrictions using different prior probability distributions. For the NiWU approach, Section 2.2.2, we use the independent prior specification and specify the inverse Wishart distribution using the two most popular parametrizations, which are either the improper prior specification or the specification by Kadiyala and Karlsson (1997).[10] For the Np(B) approach, Section 2.2.3, we introduce prior independence between $\boldsymbol{\pi}$ and B and specify $p(B)$ as discussed in Section 2.2.4, setting $\psi_1 = 0.8$ and $\psi_2 = 1.5$ for the illustration. For all models estimated, we set $\boldsymbol{\mu}_\pi = \mathbf{0}$ and $V_\pi^{-1} = 0$ for both the NiWU and the Np(B). All models include a constant term and 8 lags, as in the DGP.

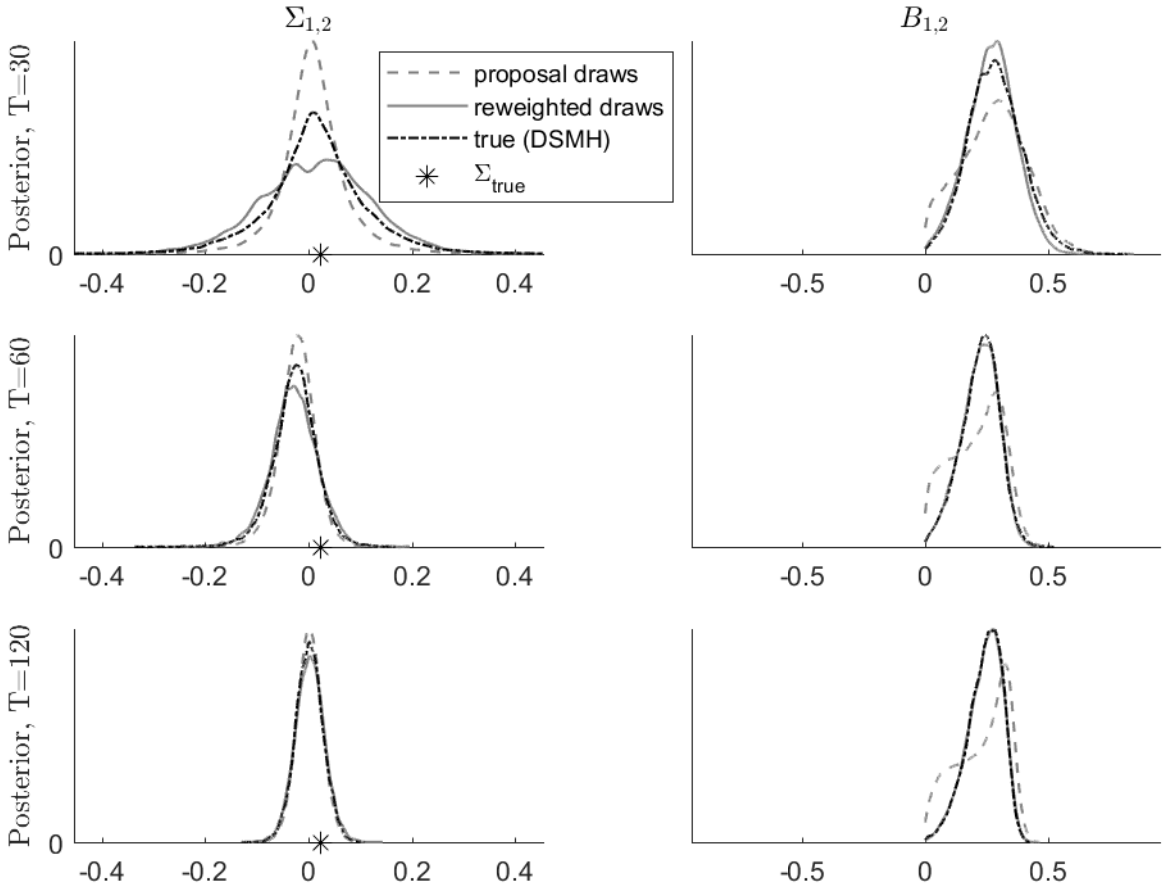
2.3.2 The intuition behind our importance sampler

We illustrate the intuition behind our posterior sampler by showing the different probability distributions involved in our algorithm. Figure 2.1 shows the results for the $(1, 2)$ entry of Σ and B for some of the datasets considered (see Section 2.A.5 of the Appendix for the full illustration). The left column of Figure 2.1 shows the results for Σ and displays the marginal distributions of $\Sigma_{1,2}$ associated with $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$. These are the importance density and the target density in Stage A of the algorithm, respectively. $p(\Sigma|Y)_{Np(B)}$ is sampled using both our algorithm and the Dynamic Striated Metropolis-Hastings algorithm by Waggoner et al. (2016). The closer these two empirical distributions are, the more the algorithm successfully explores the posterior distribution of interest. The right column of Figure 2.1 reports the equivalent distributions for B. It shows the marginal distribution of $p(B_{1,2}|Y)_{Np(B)}$ explored using either our algorithm or the Dynamic Striated Metropolis-Hastings (DSMH) algorithm, and the proposal distribution obtained when mapping draws from $p(\Sigma|Y)_{Np(B)}$ into B using draws from $p(Q|\Sigma)_{NiWU}$. See Table 2.5 in the Appendix for how we set the tuning parameters required in our algorithm, and Table 2.7 for the diagnostics on the importance weights.

As we see from the left column of Figure 2.1, the dataset with $T = 30$ observations is still too small for $p(\Sigma|Y)_{NiWU}$ (dashed line) to be similar to $p(\Sigma|Y)_{Np(B)}$ (sampled by the DSMH, dotted-dashed line), making the reweighted draws a poor approxima-

[10] The improper prior specification sets $d = 0$ and $S = 0 \cdot I_k$. The parametrization by Kadiyala and Karlsson (1997) sets $d = k + 2$, and sets S such that $E(\Sigma)$ equals the diagonal matrix displaying, on the diagonal, the variance of the residuals in univariate autoregressive processes, estimated on a training sample.

Figure 2.1: Illustration of our algorithm



Note: In the left column, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the right column, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$. See Figure 2.7 to Figure 2.11 in the Appendix for the full illustration.

tion of $p(\Sigma|Y)_{Np(B)}$. However, as the sample size increases, the likelihood dominates. This makes $p(\Sigma|Y)_{NiWU}$ an excellent importance function for $p(\Sigma|Y)_{Np(B)}$ already for $T = 60$, such that the reweighted draws now well approximate $p(\Sigma|Y)_{Np(B)}$ from the DSMH sampler. The right column of the figure displays how successful our algorithm is in sampling the posterior $p(B|Y)_{Np(B)}$. While for $T = 30$ the distribution $p(\Sigma|Y)_{Np(B)}$ is still quite different from the distribution generating proposal draws, the associated sampling of $p(B|Y)_{Np(B)}$ is already close to what is detected by the Dynamic Striated Metropolis-Hastings algorithm. For the remaining datasets the approximation improves even further, showing that our sampler is successful in sampling the posterior distribution of interest.

Table 2.1: Performance of our algorithm

T	$T-p$	Stage A			Stage B		
		proposal draws	effective sample size	relative effective sample size	proposal draws	effective sample size	relative effective sample size
		m_2	ESS^A	$\frac{ESS^A}{m_2}$	m_5	ESS^B	$\frac{ESS^B}{m_5}$
30	22	50,000	1,894	0.0379	80,000	45,322	0.5665
60	52	50,000	27,594	0.5519	80,000	56,260	0.7033
120	112	50,000	38,853	0.7771	80,000	53,740	0.6718
240	232	50,000	44,156	0.8831	80,000	52,522	0.6565
480	472	50,000	46,672	0.9334	80,000	51,858	0.6482

Table 2.1 provides additional intuition for our algorithm by reporting relevant metrics from the sampler. As should be expected, the higher is the sample size of the dataset, the higher is the effective sample size in Stage A, further confirming that the importance function tends to coincide with the target distribution. By contrast, in Stage B a high effective sample size is not required for the sampler to successfully explore the posterior distribution. As Table 2.1 shows, approximately 55% of the initial draws are effectively used already for $T = 60$, a ratio that does not change much as the sample size increases.

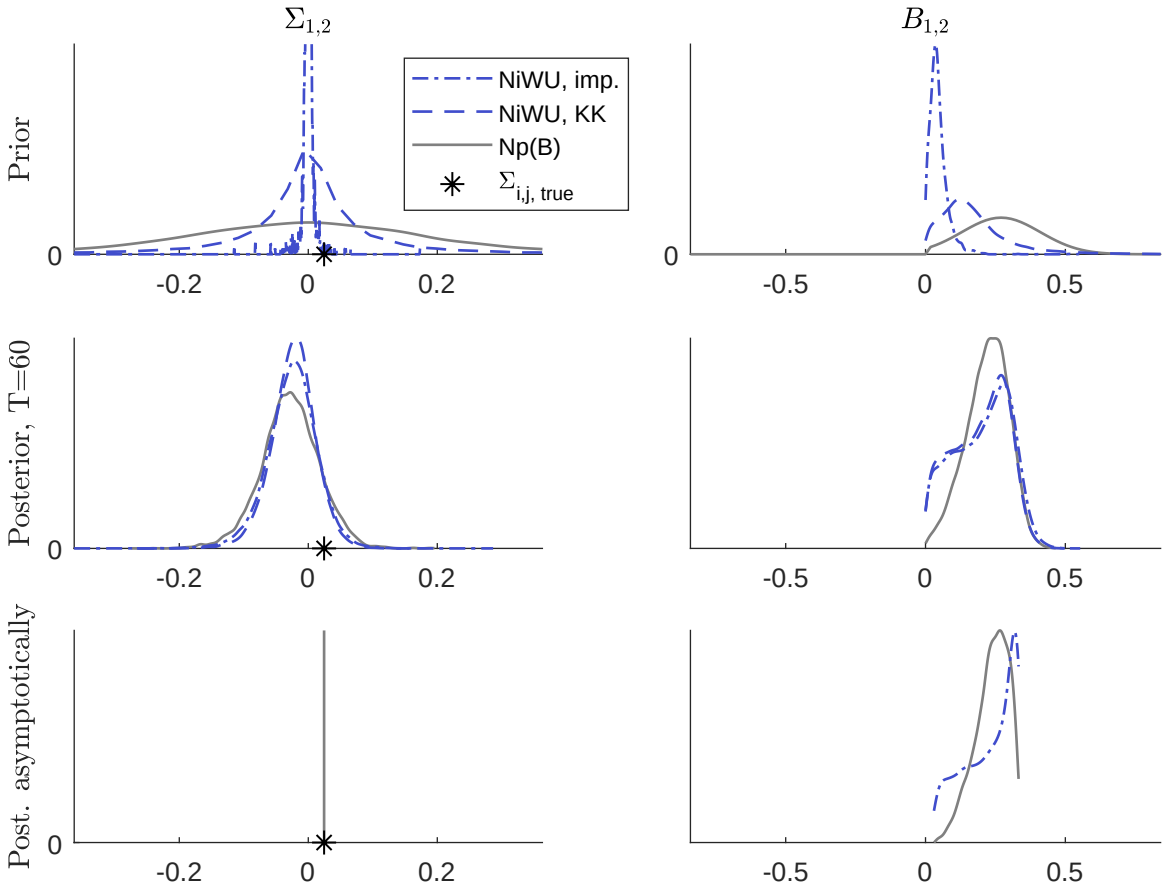
2.3.3 Comparison to the NiWU approach

Having discussed the key intuition of the sampler, we now illustrate what drives the difference between the Np(B) and the NiWU approach, and compare the computational time. To improve the comparison, for each dataset we run the NiWU approach to generate the same number of draws that are effectively obtained from the Np(B).

Figure 2.2 shows the equivalent of Figure 2.1 by reporting prior and posterior distributions associated with our Np(B) approach and with the traditional NiWU approach.[11] As we see from Figure 2.2, the prior distributions on $\Sigma_{1,2}$ are quite different, but the associated posterior distributions are very similar already for $T = 60$. By contrast, since B is not identified, differences in prior beliefs on B between the NiWU and the Np(B) approach remain present in the posterior distributions also in a large sample, as $p(Q|\Sigma)_{NiWU}$ and $p(Q|\Sigma)_{Np(B)}$ differ. Figure 2.2 also shows that the posterior distributions associated with the two parametrizations of the NiWU approach are quite similar already for $T = 60$. This occurs because, upon learning from the data about Σ, the remaining posterior uncertainty on B largely comes from $p(Q|\Sigma)_{NiWU}$, which is the same irrespectively of the parametrization of the inverse Wishart prior. Note also that $p(B|Y)_{Np(B)}$ is tighter than $p(B|Y)_{NiWU}$ despite $p(B)_{Np(B)}$ being wider

[11] The prior distribution in the NiWU case with the improper prior specification is approximated using $d = k + 2$ and $S = 0.01 \cdot I_2$, see Figure 2.13 in the Appendix.

Figure 2.2: Comparison to the NiWU approach

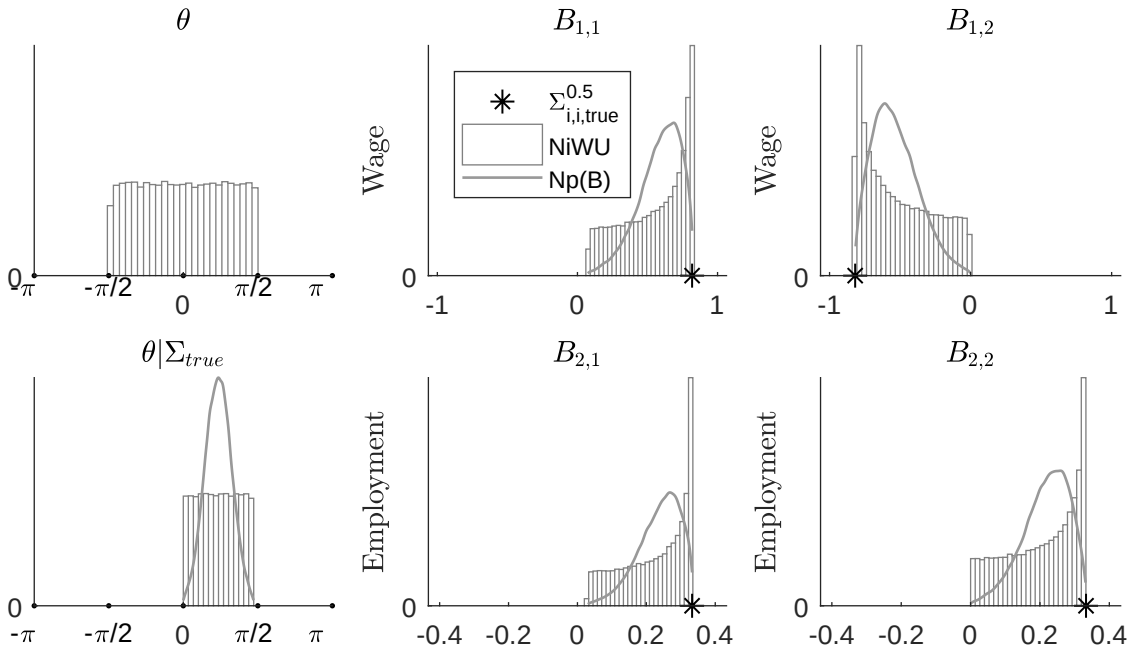


Note: See Figure 2.12 to Figure 2.19 in the Appendix for the full illustration.

than $p(B)_{NiWU}$. This happens because $p(B|Y)_{NiWU}$ inherits posterior uncertainty from $p(Q|\Sigma)_{NiWU}$, which does not take an explicit stand on which part of the structural parameter space the researcher considers more reasonable, but accepts what is implied by the uniformity on $\mathcal{Q}_\Sigma$.

The intuition behind the differences between the Np(B) and the NiWU approach can be further clarified by abstracting from estimation uncertainty and comparing $p(Q|\Sigma)_{NiWU}$ to $p(Q|\Sigma)_{Np(B)}$. In the bivariate case, distributions on Q can be shown graphically as the distribution on the corresponding rotation angle θ of Givens transformations matrices (see, for example, Fry and Pagan, 2011, as well as the analysis in Baumeister and Hamilton, 2015). Uniformity on Q is equivalent to uniformity on θ. The top-left plot of Figure 2.3 shows that indeed the angle of the rotation matrices that replicate draws of Q from the algorithm by Rubio-Ramirez et al. (2010) is uniformly distributed in the support $[-\pi/2, \pi/2]$. Conditioning on Σ_{true}, the rotation angles consistent with the sign restrictions are the subset shown in the bottom-left

Figure 2.3: Rotation angle implicit in $p(Q|\Sigma)_{NiWU}$ and $p(Q|\Sigma)_{Np(B)}$



Note: The figure shows the distribution of the rotation angle that ensures $\tilde{Q} = Q(\theta)$, with $Q(\theta)$ the Givens transformations matrix $\begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$ and Q a draw from either $p(Q|\Sigma)_{NiWU}$ or $p(Q|\Sigma)_{Np(B)}$. See Section 2.A.3.2 of the Appendix for further details.

plot of the figure, which correspond to $\mathcal{Q}_{\Sigma_{true}}$. While the NiWU approach treats such angles as equally plausible, the Np(B) approach does not, preferring instead to take an explicit stand on the part of the structural parameter space that is considered more in line with the scaling of the variables. The remaining panels of Figure 2.3 show the implied distribution on B. Given the constraint from equation (2.11), no draw of b_{ij} is obtained outside of the interval $[-\Sigma^{0.5}_{i,i,true}, +\Sigma^{0.5}_{i,i,true}]$, as displayed in the figure. The NiWU approach implies a distribution that is skewed towards such bounds (equation (33) in Baumeister and Hamilton, 2015 and their Figure 1), while the Np(B) approach implies a distribution that reflects $p(B)_{Np(B)}$. As the sample size increases, the posterior distributions $p(B|Y)_{NiWU}$ and $p(B|Y)_{Np(B)}$ approach the ones displayed in Figure 2.3.[12]

[12]Figure 2.3 shows the analysis conditioning on Σ_{true}. As the sample size increases, both $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$ collapse to a point mass at Σ_{true}, making the analysis conditioning on Σ_{true} relevant as a discussion of the posterior distributions $p(B|Y)_{NiWU}$ and $p(B|Y)_{Np(B)}$. Within the NiWU approach, the fact that the prior beliefs on B differ across parametrizations while still leading to almost identical posteriors suggests that it can be misleading to inspect prior beliefs on structural parameters to study what information the NiWU approach introduces on the results. It is, instead, best to consider the analysis conditioning on Σ_{true}. Note also that the uniform distribution in the full space $\mathcal{Q}$ explored by the algorithm by Rubio-Ramirez et al. (2010) is used in our algo-

Table 2.2: Comparison of the computational time

	Np(B) approach												**NiWU approach**					
	Our algorithm									DSMH algorithm			with improper prior			with KK(1997) prior		
	Stage A			Stage B			Total											
T	h	m	s	h	m	s	h	m	s	h	m	s	h	m	s	h	m	s
30		5	52			20		6	12		3	40			24			22
60		5	32			17		5	50		4	35			30			29
120		5	34			17		5	52		7	5			32			31
240		5	37			17		5	55		28	8			34			32
480		5	52			17		6	9	2	25	12			37			37

Note: All codes are run on Matlab, except for the Dynamic Striated Metropolis-Hastings algorithm, which we coded on Fortran to reduce computational time.

Table 2.2 shows the computational time of the Np(B) approach, the NiWU approach, and of the Dynamic Striated Metropolis-Hastings algorithm. All applications of the Np(B) approach and the NiWU approach take only a few minutes to run on Matlab. The Dynamic Striated Metropolis-Hastings algorithm takes longer to run due to its sequential nature.

2.4 Application to the oil market

We now apply our methodology to real data and revisit the model of the oil market by Kilian and Murphy (2012). We show that inference becomes sharper when taking into account the scaling of the variables in forming prior beliefs to introduce the same sign restrictions as Kilian and Murphy (2012). The exercise also illustrates to what extent the results are affected by the actual prior probability distribution used to express the sign restrictions.

2.4.1 The model

We use the three-variate model by Kilian (2009) and Kilian and Murphy (2012), which has become standard in the literature. The model includes the percentage variation in global crude oil production, the detrended index of global real economic activity developed by Kilian (2009), and the log of the real price of oil, multiplied by 100. We use the data updated by Antolín-Díaz and Rubio-Ramírez (2018), which covers

rithm as a devise to explore the distribution that is uniform in $\mathcal{Q}_{\Sigma_{true}}$. The marginal distributions $p(Q)_{NiWU} = \int p(Q|\Sigma)_{NiWU}\, p(\Sigma)_{NiWU}\, d\Sigma$ and $p(Q|Y)_{NiWU} = \int p(Q|\Sigma)_{NiWU}\, p(\Sigma|Y)_{NiWU}\, d\Sigma$ are not necessarily uniform nor they have mass in the full space $\mathcal{Q}$ (Figure 2.22 in the Appendix) .

Table 2.3: Sign restrictions on the contemporaneous impulse responses

A) Sign restrictions used

	oil supply shock	aggregate demand shock	oil-specific demand shock
Oil production	−	+	+
economic activity	−	+	−
real price of oil	+	+	+

B) Prior distributions modelling the sign restrictions

		ψ_1	ψ_2	$Prob(\lvert b_{i,j}\rvert > \gamma_i)$
prior I	*wide prior*	2	4	0.83
prior II	*medium prior*	1	2	0.53
prior III	*tight prior*	0.8	1.5	0.33

the period from January 1971 to December 2015. To improve the comparability with Antolín-Díaz and Rubio-Ramírez (2018) we add a constant and 24 lags in the model. We use the independent NiWU specification and set $\boldsymbol{\mu}_\pi = \mathbf{0}$ and $V_\pi^{-1} = 0$ for both the NiWU approach and our Np(B).

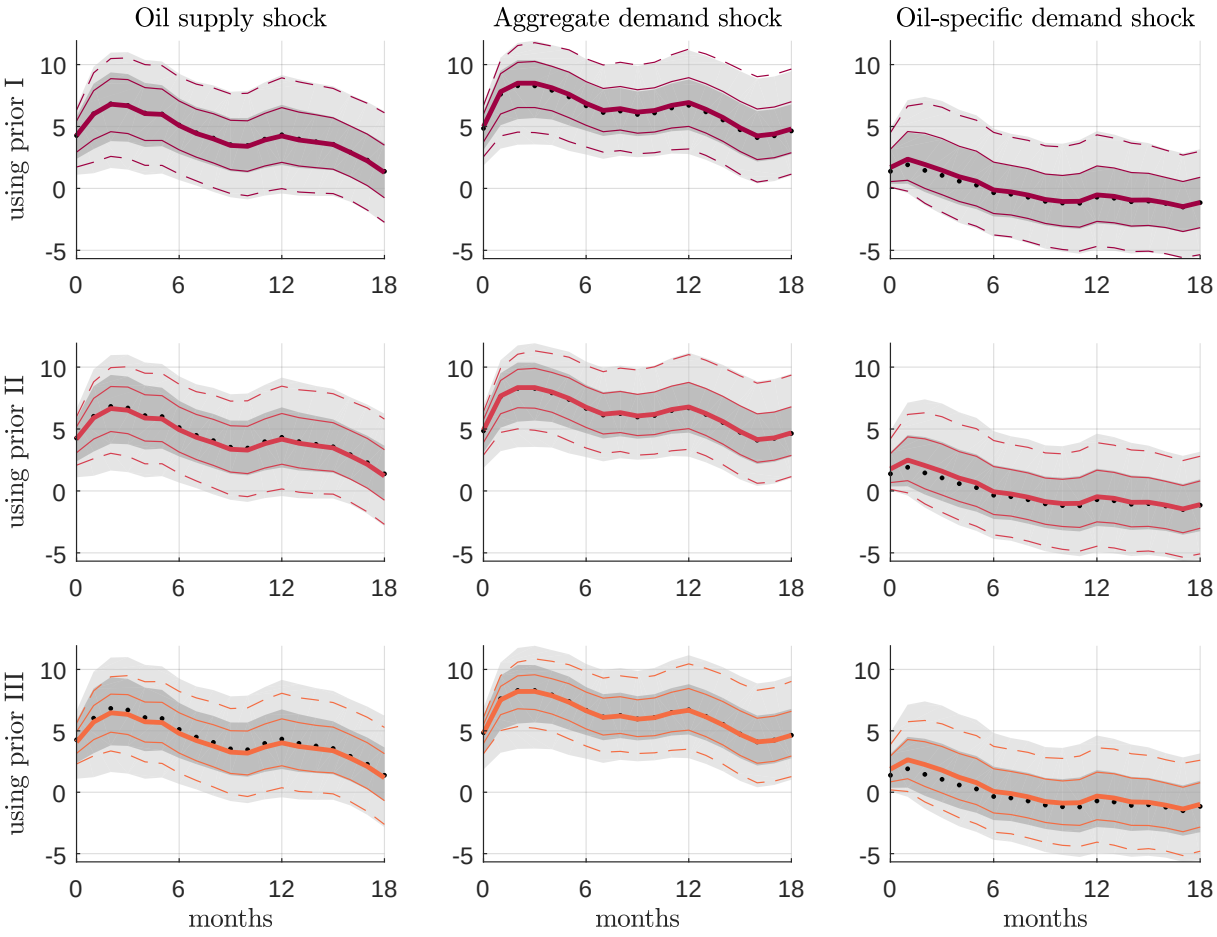
We label the structural shocks using the sign restrictions on the contemporaneous impulse responses employed by both Kilian and Murphy (2012) and Antolín-Díaz and Rubio-Ramírez (2018), see Table 2.3. However, we depart from both papers along two dimensions. First, we do not introduce explicit restrictions on elasticities, nor on the historical decompositions. Second we do not model the sign restrictions through the NiWU approach, but through the prior distribution proposed in Section 2.2.4. As discussed, this prior first uses a training sample to estimate an indicative upper bound γ_i for the elements $b_{i,j}$, and then allocates prior mass by selecting the hyperparameters ψ_1 and ψ_2. ψ_1 affects the first moment of the marginal prior distribution in $p(b_{i,j})$, whose mode is set equal to $\pm\psi_1\gamma_i$, while ψ_2 controls for the second moment of the prior by being positively related to the prior mass allocated to $|b_{i,j}| \geq \gamma_i$. We explore the role of prior beliefs by using the three separate specifications for ψ_1 and ψ_2 documented in Table 2.3. Prior I corresponds to a wide prior that attaches approximately 80% probability mass beyond the estimated γ_i. Priors II and III progressively tighten the prior to make it more consistent with the scaling of the variables, giving approximately 50% and 30% prior probability mass beyond γ_i, respectively. We favour Prior III, which implies a mode of the marginal prior slightly below the estimated upper bound, while still allowing for a non-negligible tail that gives prior mass above this point.[13]

[13]We estimate the scale γ_i using a training sample on the first 20% of the available observations, as in Primiceri (2005). The prior distributions, which are shown in Figure 2.26 of the Appendix, are such that the marginal prior on the effect of different shocks on each variable only differ potentially up to sign but not magnitude, in order not to introduce asymmetries in the results. See also Table 2.13 for a further illustration of the distribution of the probability mass under the prior distribution.

2.4.2 Results

To make the analysis more focused, we concentrate our discussion on the drivers of oil price variations, comparing the results from our Np(B) approach to the results from the NiWU approach, parametrized with the independent improper prior specification. We refer to Section 2.A.6 of the Appendix for the analysis of the other variables in the model, as well as for robustness checks and diagnostics on the performance of the sampler.

Figure 2.4: Posterior impulse responses for the real oil price, comparing NiWU and Np(B)



Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. The rows of the figure differ for the parametrization used for the prior distribution $p(B)$, as from Table 2.3. See Figure 2.34 to Figure 2.37 in the Appendix for the full analysis.

Figure 2.4 shows how one-standard-deviation shocks affect the price of oil, and compares the results from the NiWU approach and from our Np(B) approach. The columns

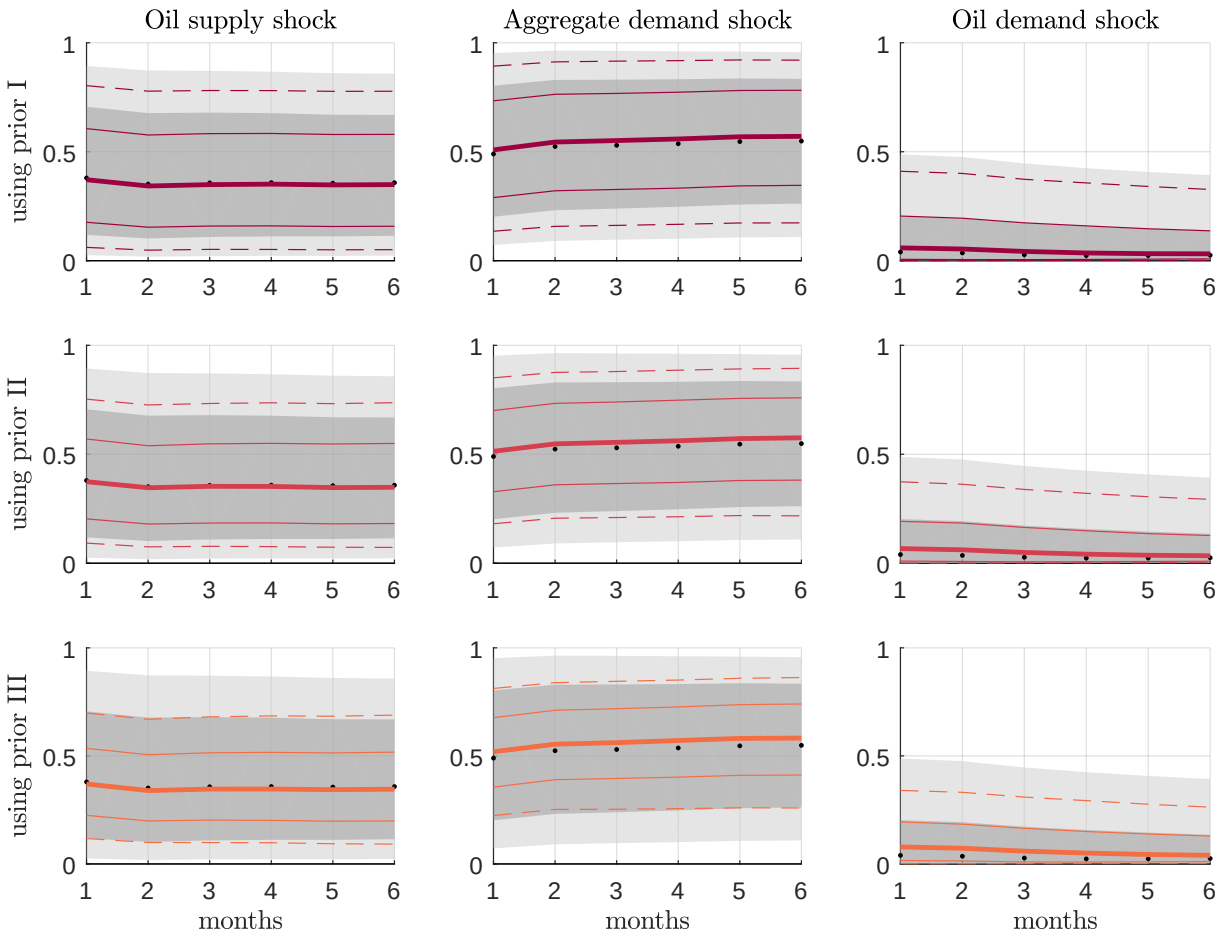
of the figure differ for the structural shock considered, while the rows differ for whether prior specification I, II or III is used in the Np(B) approach. The figure shows that the pointwise posterior bands associated with the NiWU approach are quite wide. Indeed, it is this feature that led Kilian and Murphy (2012) and Antolín-Díaz and Rubio-Ramírez (2018) to introduce further restrictions on elasticities and/or shocks and historical decompositions. The posterior bands associated with the Np(B) approach can, instead, be tighter, depending on the prior specification used. We find that, irrespectively of the prior specification, the dataset is sufficiently large for the NiWU and the Np(B) approach to deliver nearly identical posteriors for the bounds $\pm\Sigma_{i,i}^{0.5}$ that constrain $b_{i,j}$ through equation (2.11) (see Figure 2.25 and Figure 2.31-Figure 2.33 in the Appendix). Accordingly, the differences in $p(B|Y)_{NiWU}$ and $p(B|Y)_{Np(B)}$ are strongly influenced by differences in $p(Q|\Sigma)_{NiWU}$ (which is the uniform distribution in $\mathcal{Q}_\Sigma$) and $p(Q|\Sigma)_{Np(B)}$ (which is the distribution implied by the prior beliefs $p(B)_{Np(B)}$ used).

As shown in the first row of Figure 2.4, prior I from our Np(B) approach replicates the posterior bands from the NiWU approach up to a close approximation. Yet, prior I attaches as much as 80% prior mass to values of B above the estimated reasonable bound γ_i. We view this prior mass as too wide given the scaling of the variables. As shown with prior II and III, tightening $p(B)$ to make it more consistent with the scaling of the variables tightens the posterior bands considerably. On the short horizon of the response, the 95% credible bands associated with the Np(B) approach under prior III are as tight as the 68% credible bands of the NiWU approach. This suggests that introducing explicit information on the scaling of the variables can make inference much sharper. Using prior III, which attaches approximately 30% prior mass to values of B above the estimated upper bounds γ_i, we find that oil-specific demand shocks generate an immediate increase in the price of oil, an increase that then progressively declines, while aggregate demand shocks produce stronger effects also at longer horizons. While this confirms the results by Kilian (2009) that demand shocks are important drivers of oil price responses, we find that this is more so for aggregate demand shocks rather than oil-specific demand shocks. In addition, we find that oil supply shocks generate sizeable effects on the price of oil, although with smaller effects when focusing on the longer horizon of the response.

The result on the importance of oil supply shocks in driving the price of oil is in line with the results by Caldara et al. (2018) and Baumeister and Hamilton (2019) despite the different methodologies used. Caldara et al. (2018) use an exactly identified model that minimizes the distance between the elasticities implied by the VAR model and external estimates. Yet, as they show, the parametrization of the elasticities have an important effect on the results. Baumeister and Hamilton (2019) also build their

analysis on external information on price elasticities on oil, and use a sign restricted framework. They then add information on the dynamics in inventories and measurement error, weight data differently depending on the period that they correspond to, and combine sign restrictions on elasticities with sign restrictions on the contemporaneous impulse responses. We show that the results in Caldara et al. (2018) and Baumeister and Hamilton (2019) are robust to a framework that focuses on the sign restrictions on the contemporaneous impulse responses. Figure 2.48 in the Appendix shows that the posterior distributions on the price elasticities implicit in our approach are broadly consistent with the estimates by Caldara et al. (2018) and Baumeister and Hamilton (2019).

Figure 2.5: Posterior forecast error variance decomposition for the real oil price, comparing NiWU and Np(B)

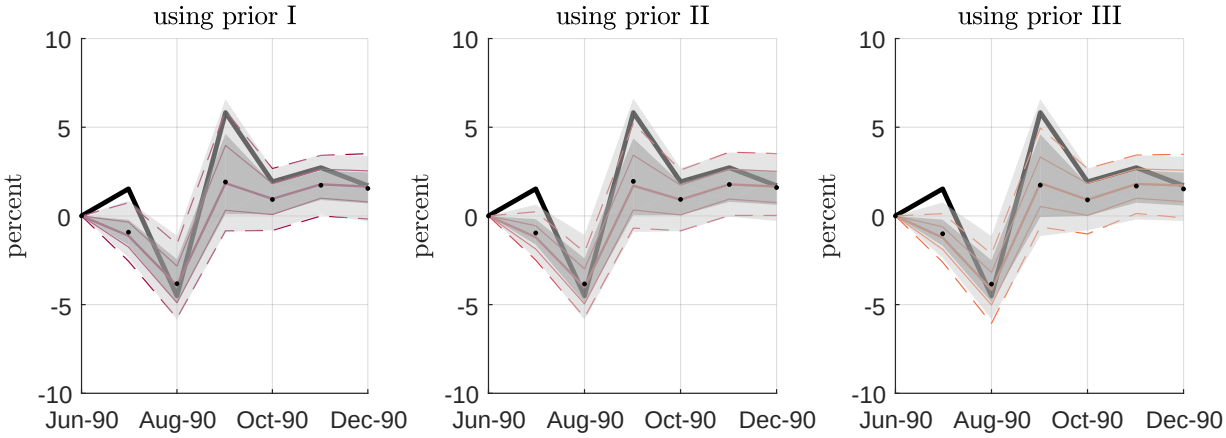


Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. The rows of the figure differ for the parametrization used for the prior distribution $p(B)_{Np(B)}$, as from Table 2.3. See Figure 2.38 to Figure 2.40 in the Appendix for the full analysis.

The analysis of forecast error variance decompositions, displayed in Figure 2.5, shows that the NiWU approach can deliver credible bands that are too wide to imply results that can be interpreted. The 95% pointwise credible band can go from close to 0 to close to 1, failing to disclose the role of the structural shocks in driving the variance of forecast errors. By contrast, inference is much sharper when prior mass on key structural parameters is ensured to be in line with the scaling of the variables. As we move our prior from specification I to III, we find that the unexpected variations in the price of oil are mainly driven by supply shocks and aggregate demand shocks for approximately 20-50% and 30-60%, respectively, while oil-specific demand shocks have a more subdued effect. The result that supply shocks have an important role in driving unexpected variations in the price of oil is consistent with Caldara et al. (2018).

Figure 2.6: Historical decomposition, cumulative effects of the shocks

1990: oil supply shocks → oil production growth



Note: The figure shows the data (solid black line) and its decomposition into the cumulative contribution of the estimated structural shocks from the beginning of the sample until period t. The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. Having subtracted the value corresponding to June 1990 before computing pointwise statistics, the figure can be interpreted as percent relative to the initial point. See Figure 2.45 to Figure 2.47 in the Appendix for the full analysis.

We conclude the analysis by further relating our work to Antolín-Díaz and Rubio-Ramírez (2018). Antolín-Díaz and Rubio-Ramírez (2018) achieve a sharpening of the posterior credible sets by introducing the restriction that oil supply shocks matter significantly in driving the drop in oil production in August 1990. Indeed, this is the key event in their application, as they discuss. Figure 2.6 shows that our approach delivers this feature as a result, rather than as a restriction. As we make our prior more in line with the scale of the variables, the credible sets associated with our methodology leave

little doubt that oil supply shocks were relevant drivers of the drop in the oil production. By contrast, the NiWU approach delivers wide posterior bands, leading Antolín-Díaz and Rubio-Ramírez (2018) to introduce the restriction. The result that supply shocks contributed to the decline in oil production in August 1990 is also reported by Caldara et al. (2018).

2.5 Conclusions

Structural Vector Autoregressive models are frequently identified using sign restrictions on the impulse response of selected structural shocks of interest. However, it is not clear how this identification approach should be implemented in practice. On the one hand, it is convenient to start from a specification on reduced form parameters, as this makes posterior sampling highly tractable. On the other hand it is important to retain flexibility on the prior beliefs implied for the key structural parameters of interest, since such prior affects the posterior distribution even in a large sample.

We propose an approach that offers flexibility for the prior specification on the impulse response horizon that matters the most, while ensuring that the joint posterior distribution is tractable. We illustrate the intuition of our approach using simulations on the bivariate demand and supply model by Baumeister and Hamilton (2015). We then develop an application to the oil market and show that our approach delivers sharper inference. Consistent with Baumeister and Hamilton (2019) and Caldara et al. (2018), we find that oil supply shocks have a comparable role in explaining oil price dynamics relative to oil-specific demand shocks.

2.A Further Results

2.A.1 Likelihood function of the model

To derive the likelihood of the model, we start from equation (2.1) of the paper, which we rewrite here for convenience:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_t + B\boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I_k). \tag{2.12}$$

$\boldsymbol{y}_t$ is a $k \times 1$ vector of variables, $\boldsymbol{\epsilon}_t$ is a $k \times 1$ vector of structural shocks, $\boldsymbol{w}_t$ is an $m \times 1$ vector of lagged variables and potentially deterministic controls, Π is a $k \times m$ matrix of reduced form parameters, and B is a $k \times k$ matrix of structural parameters. Write the model in compact form as

$$Y = \Pi W + BE, \tag{2.13}$$

where $Y = [\boldsymbol{y}_1, ..., \boldsymbol{y}_t, ..., \boldsymbol{y}_T]$ and $E = [\boldsymbol{\epsilon}_1, ..., \boldsymbol{\epsilon}_t, ..., \boldsymbol{\epsilon}_T]$ are $k \times T$ matrices of data and shocks, and $W = [\boldsymbol{w}_1, ..., \boldsymbol{w}_t, ..., \boldsymbol{w}_T]$ is an $m \times T$ matrix of data. Then, make use of the formula $vec(\bar{A}\bar{B}\bar{C}) = (\bar{C}' \otimes \bar{A}) \cdot vec(\bar{B})$ (see Lütkepohl, 2005, mathematical appendix) and rewrite the model as

$$\tilde{\boldsymbol{y}} = Z\boldsymbol{\pi} + (I_T \otimes B)\tilde{\boldsymbol{\epsilon}}, \qquad \tilde{\boldsymbol{\epsilon}} \sim N\Big(\mathbf{0}, (I_T \otimes I_k)\Big), \tag{2.14}$$

with $\tilde{\boldsymbol{y}} = vec(Y)$ and $\tilde{\boldsymbol{\epsilon}} = vec(E)$ of dimension $kT \times 1$ and $Z = (W' \otimes I_k)$ of dimension $kT \times mk$. The $mk \times 1$ vector $\boldsymbol{\pi} = vec(\Pi)$ stacks the columns of Π vertically. Last, rewrite the model in reduced form, obtaining

$$\tilde{\boldsymbol{y}} = Z\boldsymbol{\pi} + \tilde{\boldsymbol{u}}, \qquad \tilde{\boldsymbol{u}} \sim N\Big(\mathbf{0}, \tilde{\Sigma}\Big), \tag{2.15}$$

with $\tilde{\boldsymbol{u}} = (I_T \otimes B)\tilde{\boldsymbol{\epsilon}}$. $\tilde{\Sigma}$ and B are related through the equality

$$\tilde{\Sigma} = \big(I_T \otimes BB'\big). \tag{2.16}$$

The likelihood function can now be written in B and $\boldsymbol{\pi}$ as

$$p(Y|\boldsymbol{\pi}, B) = (2\pi)^{-\frac{kT}{2}} |det(I_T \otimes BB')|^{-\frac{1}{2}} e^{-\frac{1}{2}(\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi})'(I_T \otimes BB')^{-1}(\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi})}, \tag{2.17}$$

$$= (2\pi)^{-\frac{kT}{2}} |det(BB')|^{-\frac{T}{2}} e^{-\frac{1}{2}(\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi})'(I_T \otimes BB')^{-1}(\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi})}. \tag{2.18}$$

2.A.2 NiWU approach used in the literature

In this section we report the key equations related to the NiWU approach and outline the corresponding algorithm. We limit the illustration to the independent (rather than the conjugate) Normal-inverse-Wishart-Uniform prior specification for two reasons. First, because it is more convenient for the implementation of our algorithm, as we discuss in section 2.A.3. Second, because it makes the NiWU approach more comparable with the Np(B) approach proposed in the paper, given that they can use the same prior on $\boldsymbol{\pi}$. We refer the interested reader to Canova (2007), Koop et al. (2010) and Kilian and Lütkepohl (2017) for a more thorough discussion of the NiWU approach, and to the material available on our website for the derivations.

Under the independent Normal-inverse-Wishart-Uniform approach the joint prior distribution is given by[14]

$$p(\boldsymbol{\pi}, \Sigma, Q) = p(Q|\Sigma) \cdot p(\boldsymbol{\pi}) \cdot p(\Sigma), \tag{2.20}$$

with

$$\begin{aligned} \boldsymbol{\pi} &\sim N(\boldsymbol{\mu}_{NiWU}, V_{NiWU}), && (2.21)\\ \Sigma &\sim iW(d_{NiWU}, S_{NiWU}), && (2.22)\\ Q|\Sigma &\sim U. && (2.23) \end{aligned}$$

This leads to the joint posterior distribution

$$p(\boldsymbol{\pi}, \Sigma, Q|Y) = p(Q|\Sigma) \cdot p(\boldsymbol{\pi}, \Sigma|Y), \tag{2.24}$$

with

$$\begin{aligned} \boldsymbol{\pi}|Y, \Sigma &\sim N(\boldsymbol{\mu}^*_{NiWU}, V^*_{NiWU}), && (2.25)\\ \Sigma|Y, \Pi &\sim iW(d^*_{NiWU}, S^*_{NiWU}), && (2.26)\\ Q|\Sigma &\sim U, && (2.27) \end{aligned}$$

[14]Since sign restrictions on B potentially imply that a subset of the parameter space of Σ admits no matrix B satisfying the sign restrictions, the joint prior probability distribution effectively used on reduced form parameters in the structural analysis is

$$p(\boldsymbol{\pi}, \Sigma) = I\{\Sigma\} \cdot p(\boldsymbol{\pi}) p(\Sigma), \tag{2.19}$$

with $I\{\Sigma\}$ an indicator function taking value of 1 if Σ implies a non-empty set of B satisfying the sign restrictions. We omit the additional indicator function to simplify the notation.

and

$$V^*_{NiWU} = \Big(V^{-1}_{NiWU} + \big[WW' \otimes \Sigma^{-1}\big]\Big)^{-1}, \tag{2.28}$$

$$\boldsymbol{\mu}^*_{NiWU} = V^*_{NiWU} \cdot \Big(V^{-1}_{NiWU}\boldsymbol{\mu}_{NiWU} + \big[W \otimes \Sigma^{-1}\big]\tilde{\boldsymbol{y}}\Big), \tag{2.29}$$

$$d^*_{NiWU} = d_{NiWU} + T, \tag{2.30}$$

$$S^*_{NiWU} = S_{NiWU} + (Y - \Pi W)(Y - \Pi W)'. \tag{2.31}$$

Equations (2.25) and (2.26) can be used in a Gibbs sampler to explore the joint posterior $p(\boldsymbol{\pi}, \Sigma|Y)$. The marginal posterior distribution for Σ associated with equations (2.25) and (2.26) equals[15]

$$\begin{aligned} p(\Sigma|Y) \propto & |det(\Sigma)|^{-\frac{d^*_{NiWU}+k+1}{2}} \cdot |det(V^*_{NiWU})|^{\frac{1}{2}} \cdot \\ & \cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'(I_T \otimes \Sigma^{-1})\tilde{\boldsymbol{y}} - \boldsymbol{\mu}^{*'}_{NiWU} V^{*^{-1}}_{NiWU} \boldsymbol{\mu}^*_{NiWU} + tr\left[\Sigma^{-1} S_{NiWU}\right]\right\}}, \end{aligned} \tag{2.32}$$

which is required in our algorithm.

Given a general distribution $p_\Sigma(\Sigma)$ for Σ, a general distribution $p_Q(Q|\Sigma)$ for Q, and defining B such that $B = h(\Sigma)Q$, it holds that

$$p_B(B) = J(\Sigma \to B) \cdot p_\Sigma\big(\Sigma = BB'\big) \cdot p_Q\big(Q|\Sigma = BB'\big), \tag{2.33}$$

with $J(\Sigma \to B)$ the Jacobian of the transformation from Σ to B (Bibby et al., 1979, Mathai and Haubold, 2008, Arias et al., 2018). When no zero restrictions are introduced on B,[16]

$$J(\Sigma \to B) = |det(B)|. \tag{2.34}$$

It follows that the implicit distributions on B associated with equations (2.22), (2.32) and (2.23) are

$$p(B) \propto I\{\mathrm{B}\} \cdot |det(B)|^{-(d_{NiWU}+k)} \cdot e^{-\frac{1}{2}\left\{vec(B^{-1})'(S_{NiWU}\otimes I_k)vec(B^{-1})\right\}} \tag{2.35}$$

$$\begin{aligned} p(B|Y) \propto & I\{\mathrm{B}\} \cdot |det(B)|^{-(d^*_{NiWU}+k)} \cdot |det(V^*_{NiWU})|^{\frac{1}{2}} \cdot \\ & \cdot e^{-\frac{1}{2}\left\{vec(B^{-1})'(S_{NiWU}\otimes I_k)vec(B^{-1}) + \tilde{\boldsymbol{y}}'\left(I_T \otimes (BB')^{-1}\right)\tilde{\boldsymbol{y}} - \boldsymbol{\mu}'^*_{NiWU} V^{*^{-1}}_{NiWU} \boldsymbol{\mu}^*_{NiWU}\right\}}, \end{aligned} \tag{2.36}$$

[15] Step-by-step derivations for the NiWU approach for both the independent and the conjugate prior specifications are available at https://drive.google.com/open?id=0B0CRyT66a7B2TmpWV2tjczdwM00.

[16] We compute the Jacobians of matrix transformations using the results in Bibby et al. (1979), Chapter 2, and Mathai and Haubold (2008), Chapter 11. $\Sigma = BB'$ implies $dB = |det(\Sigma)|^{-\frac{1}{2}} d\Sigma$. Substituting $|det(\Sigma)| = |det(BB')| = |det(B)|^2$ gives $J(\Sigma \to B) = \left|\frac{d\Sigma}{dB}\right| = |det(B)|$.

with $I\{\text{B}\}$ an indicator function taking value of 1 if B satisfies the sign restrictions. Equation (2.35) coincides with the first part of equation (2.8) in Arias et al. (2018), once adjusted for the difference in the notation used. We use equation (2.36) for the implementation of the Dynamic Striated Metropolis-Hastings algorithm, see section 2.A.4. If zero restrictions are introduced on B, the computation of $J(\Sigma \to B)$ requires a numerical procedure, see Arias et al. (2018).

In the paper, we use the following algorithm to study the join posterior distribution $p(\boldsymbol{\pi}, \Sigma, Q|Y)$ associated with the independent NiWU approach:

NiWU algorithm:

1. run a Gibbs sampler to explore $p(\boldsymbol{\pi}, \Sigma|Y)$ based on the distributions (2.25) and (2.26) using $n_1 + n_2$ replications, with n_1 the number of burn-in replications and n_2 the number of retained replications. Store the retained draws in $\left\{\boldsymbol{\pi}_d, \Sigma_d\right\}_{d=1}^{n_2}$;
2. randomly extract $(\boldsymbol{\pi}, \Sigma)$ from $\left\{\boldsymbol{\pi}_d, \Sigma_d\right\}_{d=1}^{n_2}$ with replacement;
3. draw an orthogonal matrix (Q_d) using the method by Rubio-Ramirez et al. (2010), which extracts uniformly from the space $\mathcal{Q}$ (the full space of orthogonal matrices), and map $(\boldsymbol{\pi}_d, \Sigma_d, Q_d)$ into the structural parameters of interest;
   - 4a. if $(\boldsymbol{\pi}_d, \Sigma_d, Q_d)$ satisfy the sign restrictions (up to ordering of the shocks), store $(\boldsymbol{\pi}_d, \Sigma_d, Q_d)$ and proceed to Step 5;
   - 4b. if $(\boldsymbol{\pi}_d, \Sigma_d, Q_d)$ do not satisfy the sign restrictions, repeat Step 3 up to n_3 times. Stop as soon as $(\boldsymbol{\pi}_d, \Sigma_d, Q_d)$ satisfies the sign restrictions and proceed to Step 5, otherwise discard $(\boldsymbol{\pi}_d, \Sigma_d)$ and move back to Step 2;
5. repeat Steps 2 to 4 until a desired number of draws n_4 is obtained.[17]

[17] The algorithm is close in spirit to the algorithms used by Uhlig (2005), Rubio-Ramirez et al. (2010) and Giacomini and Kitagawa (2015). Dismissing draws of Q in Step 4b numerically approximates the extraction from a uniform distribution in $\mathcal{Q}_\Sigma$ (the subspace of $\mathcal{Q}$ that satisfies the sign restrictions given Σ). An alternative approach is discussed in Amir-Ahmadi and Drautzburg (2018). For n_3 sufficiently large in Step 4b, the algorithm delivers draws $(\boldsymbol{\pi}, \Sigma)$ that approximate $p(\boldsymbol{\pi}, \Sigma|Y)$ from equations (2.25)-(2.26) up to an indicator function taking value of one if the reduced form draws imply a non-empty set $\mathcal{Q}_\Sigma$. By contrast, when n_3 is too small, an additional tilting of these distributions is introduced, since draws $(\boldsymbol{\pi}_1, \Sigma_1)$ and $(\boldsymbol{\pi}_2, \Sigma_2)$ such that $p(\boldsymbol{\pi}_1, \Sigma_1|Y) = p(\boldsymbol{\pi}_2, \Sigma_2|Y)$ are not necessarily retained with equal probability if $\mathcal{Q}_{\Sigma_1}$ and $\mathcal{Q}_{\Sigma_2}$ differ considerably in size. Step 3 of the algorithm changes if zero restrictions are introduced, see Binning (2013), Arias et al. (2018), and Kilian and Lütkepohl (2017), Chapter 13.

Compared to the NiWU algorithm, our algorithm for the general Np(B) approach runs a Gibbs sampler only to generate *proposal* draws, which are then reweighted in the importance sampler. In the paper, we ensure that the computational time of the NiWU approach and the Np(B) approach are comparable by first running our algorithm, and then running the NiWU algorithm such that it generates a number of draws equal to the ones effectively obtained for our algorithm, see table 2.4. For example, in the simulation exercise of section 2.3 of the paper, for $T = 120$ Stage B of our algorithm effectively generates 53,740 draws. We then run the NiWU algorithm setting $n_2 = n_4 = 53,740$.

Table 2.4: Tuning parameters used for the NiWU algorithm

n_1	burn-in draws	equal to m_1 in our algorithm
n_2	retained draws	equal to ESS^B in our algorithm
n_3	maximum number of draws of Q	equal to m_4 in our algorithm
n_4	desired number of draws	equal to ESS^B in our algorithm

2.A.3 Np(B) approach proposed in the paper

In this section we provide the derivations of the posterior distribution for the approach proposed in the paper, and discuss our importance sampler. We then discuss the diagnostic procedures that we use to ensure that the weights in the importance sampler have a finite variance.

2.A.3.1 Posterior distribution

Start from the prior distribution

$$p(\boldsymbol{\pi}, B) = p(\boldsymbol{\pi}|B) \cdot p(B), \tag{2.37}$$

with

$$p(\boldsymbol{\pi}|B) = (2\pi)^{-\frac{k}{2}}|det(V_\pi)|^{-\frac{1}{2}}e^{-\frac{1}{2}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)'V_\pi^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)}, \tag{2.38}$$

where $\boldsymbol{\mu}_\pi$ and V_π can be a function of B. The joint posterior distribution then equals

$$\begin{aligned} p(\boldsymbol{\pi}, B|Y) =& p(\boldsymbol{\pi}, B)\frac{p(Y|\boldsymbol{\pi}, B)}{p(Y)}, \\ =& p(B)\cdot \\ & \cdot (2\pi)^{-\frac{k}{2}}|det(V_\pi)|^{-\frac{1}{2}}e^{-\frac{1}{2}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)'V_\pi^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)}. \\ & \cdot (2\pi)^{-\frac{kT}{2}}|det(B)|^{-T}e^{-\frac{1}{2}(\tilde{\boldsymbol{y}}-Z\boldsymbol{\pi})'(I_T\otimes BB')^{-1}(\tilde{\boldsymbol{y}}-Z\boldsymbol{\pi})}p(Y)^{-1}, \end{aligned} \tag{2.39}$$

with $p(Y) = \int_{\boldsymbol{\pi}} \int_B p(Y|\boldsymbol{\pi}, B)p(\boldsymbol{\pi}, B)d\boldsymbol{\pi}dB$. We aim to rewrite equation (2.39) as

$$p(\boldsymbol{\pi}, B|Y) = p(\boldsymbol{\pi}|B, Y) \cdot p(B|Y), \tag{2.40}$$

and to exploit analytical results for $p(\boldsymbol{\pi}|B, Y)$. To do so, define

$$\Psi(B) = \big(I_T \otimes BB'\big), \tag{2.41}$$

and rewrite the joint posterior distribution as

$$p(\boldsymbol{\pi}, B|Y) = c \cdot e^{-\frac{1}{2}\left[(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)'V_\pi^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi)+(\tilde{\boldsymbol{y}}-Z\boldsymbol{\pi})'\Psi(B)^{-1}(\tilde{\boldsymbol{y}}-Z\boldsymbol{\pi})\right]}, \tag{2.42}$$

with c a term that includes elements which are not a function of $\boldsymbol{\pi}$. As frequently done also in the NiWU approach, factorize the terms in the exponent of (2.42) as

$$\begin{aligned}
&(\boldsymbol{\pi} - \boldsymbol{\mu}_\pi)'V_\pi^{-1}(\boldsymbol{\pi} - \boldsymbol{\mu}_\pi) + (\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi})'\Psi(B)^{-1}(\tilde{\boldsymbol{y}} - Z\boldsymbol{\pi}) = \\
&\quad = \boldsymbol{\pi}'V_\pi^{-1}\boldsymbol{\pi} - 2\boldsymbol{\pi}'V_\pi^{-1}\boldsymbol{\mu}_\pi + \boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi + \\
&\quad\quad + \tilde{\boldsymbol{y}}'\Psi(B)^{-1}\tilde{\boldsymbol{y}} - 2\boldsymbol{\pi}'Z'\Psi(B)^{-1}\tilde{\boldsymbol{y}} + \boldsymbol{\pi}'Z'\Psi(B)^{-1}Z\boldsymbol{\pi} = \\
&\quad = \boldsymbol{\pi}'[V_\pi^{-1} + Z'\Psi(B)^{-1}Z]\boldsymbol{\pi} - 2\boldsymbol{\pi}'[V_\pi^{-1}\boldsymbol{\mu}_\pi + Z'\Psi(B)^{-1}\tilde{\boldsymbol{y}}] + \\
&\quad\quad + \tilde{\boldsymbol{y}}'\Psi(B)^{-1}\tilde{\boldsymbol{y}} + \boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi = \\
&\quad = (\boldsymbol{\pi} - \boldsymbol{\mu}_\pi^*)'V_\pi^{*^{-1}}(\boldsymbol{\pi} - \boldsymbol{\mu}_\pi^*) + \tilde{\boldsymbol{y}}'\Psi(B)^{-1}\tilde{\boldsymbol{y}} + \boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi - \boldsymbol{\mu}_\pi^{*'}V_\pi^{*^{-1}}\boldsymbol{\mu}_\pi^*,
\end{aligned} \tag{2.43}$$

with

$$V_\pi^* = [V_\pi^{-1} + Z'\Psi(B)^{-1}Z]^{-1}, \tag{2.44}$$

$$= \Big[V_\pi^{-1} + \big[WW' \otimes (BB')^{-1}\big]\Big]^{-1}, \tag{2.45}$$

$$\boldsymbol{\mu}_\pi^* = V_\pi^* \cdot [V_\pi^{-1}\boldsymbol{\mu}_\pi + Z'\Psi(B)^{-1}\tilde{\boldsymbol{y}}], \tag{2.46}$$

$$= V_\pi^* \cdot \Big[V_\pi^{-1}\boldsymbol{\mu}_\pi + \big[W \otimes (BB')^{-1}\big]\tilde{\boldsymbol{y}}\Big]. \tag{2.47}$$

The joint posterior distribution can now be written as

$$\begin{aligned}
p(\boldsymbol{\pi}, B|Y) = &\underbrace{(2\pi)^{-\frac{k}{2}} \cdot |det(V_\pi^*)|^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi^*)'V_\pi^{*^{-1}}(\boldsymbol{\pi}-\boldsymbol{\mu}_\pi^*)}}_{p(\boldsymbol{\pi}|B,Y)} \cdot \\
&\cdot p(B) \cdot |det(B)|^{-T} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot |det(V_\pi)|^{-\frac{1}{2}} \cdot \\
&\cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T\otimes(BB')^{-1}\right)\tilde{\boldsymbol{y}}-\boldsymbol{\mu}_\pi^{*'}V_\pi^{*^{-1}}\boldsymbol{\mu}_\pi^*+\boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi\right\}} \cdot \\
&\cdot (2\pi)^{-\frac{kT}{2}} \cdot p(Y)^{-1}.
\end{aligned}$$

It follows that

$$\boldsymbol{\pi}|B,Y \ \sim N(\boldsymbol{\mu}_\pi^*, V_\pi^*), \tag{2.48}$$

$$p(B|Y) \ \propto \ p(B)\cdot ... \tag{2.49}$$
$$\cdot\,|det(B)|^{-T}\cdot|det(V_\pi)|^{-\frac{1}{2}}\cdot|det(V_\pi^*)|^{\frac{1}{2}}\cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T\otimes(BB')^{-1}\right)\tilde{\boldsymbol{y}}-\boldsymbol{\mu}_\pi^{*\prime}V_\pi^{*-1}\boldsymbol{\mu}_\pi^*+\boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi\right\}},$$

which proves equations (2.7) and (2.8) in section 2.2.3 of the paper.

Our algorithm requires to evaluate the distribution on Σ implied by the distribution (2.49). To illustrate how to compute this function, call $\mathcal{B}(\Sigma)$ the subset of the parameter space of B such that $BB' = \bar{\Sigma}$ (note that the sets $\mathcal{B}(\Sigma)$ and $\mathcal{B}(\tilde{\Sigma})$ do not intersect unless $\bar{\Sigma} = \tilde{\Sigma}$). Then, defining $p_B(B)$ a general distribution for B, $p_\Sigma(\Sigma)$ the distribution for Σ implied by $p_B(B)$, and $\bar{\Sigma}$ a value for Σ, it holds that draws $\Sigma = \bar{\Sigma}$ are generated by $p_\Sigma(\Sigma)$ if and only if $p_B(B)$ draws from the set $\mathcal{B}(\bar{\Sigma})$. Accordingly, it holds that

$$p_\Sigma(\Sigma) = J(B\to\Sigma)\cdot\int_{\mathcal{B}(\Sigma)} p_B(B)dB, \tag{2.50}$$

where $J(B\to\Sigma)$ is the Jacobian transformation from B to Σ (which equals $|det(\Sigma)|^{-\frac{1}{2}}$ when no zero restrictions are introduced, see section 2.A.2), and $\int_{\mathcal{B}(\Sigma)} p(B)dB$ is the probability mass of $p_B(B)$ in the subspace $\mathcal{B}(\Sigma)$. In the application of the paper, this implies

$$\begin{aligned}
p(\Sigma|Y) &\propto J(B\to\Sigma)\cdot\int_{\mathcal{B}(\Sigma)} p(B|Y)dB,\\
&\propto |det(\Sigma)|^{-\frac{1}{2}}\cdot\Big[\int_{\mathcal{B}(\Sigma)} p(B)\cdot|det(B)|^{-T}\cdot|det(V_\pi)|^{-\frac{1}{2}}\cdot|det(V_\pi^*)|^{\frac{1}{2}}\cdot\\
&\qquad\cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T\otimes(BB')^{-1}\right)\tilde{\boldsymbol{y}}-\boldsymbol{\mu}_\pi^{*\prime}V_\pi^{*-1}\boldsymbol{\mu}_\pi^*+\boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi\right\}}dB\Big],\\
&\propto |det(\Sigma)|^{-\frac{T+1}{2}}\cdot|det(V_\pi)|^{-\frac{1}{2}}\cdot|det(V_\pi^*)|^{\frac{1}{2}}\cdot\\
&\qquad\cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T\otimes\Sigma^{-1}\right)\tilde{\boldsymbol{y}}-\boldsymbol{\mu}_\pi^{*\prime}V_\pi^{*-1}\boldsymbol{\mu}_\pi^*+\boldsymbol{\mu}_\pi'V_\pi^{-1}\boldsymbol{\mu}_\pi\right\}}\cdot\int_{\mathcal{B}(\Sigma)} p(B)dB.
\end{aligned} \tag{2.51}$$

Note that equation (2.51) does not necessarily coincide with equation (2.32) unless $p(B)$ coincides with the distribution implied by the NiWU approach, equation (2.36). Last, Step B of our algorithm requires evaluating the distribution $p(Q|\Sigma)$ implicit in the distribution $p(B|Y)$ (or identically, implicit in the distribution $p(B)$) within the set $\mathcal{B}(\Sigma)$. For this, note that for a general value $\tilde{B}$ with $\tilde{B}\tilde{B}' = \tilde{\Sigma}$ and $\tilde{Q} = \tilde{B}h(\tilde{\Sigma})^{-1}$,

when no zero restrictions are introduced equation (2.33) implies

$$\frac{p(B=\tilde{B})}{p(Q=\tilde{Q}|\tilde{\Sigma})} \propto J(\Sigma \to B) \cdot p(\Sigma=\tilde{\Sigma}), \tag{2.52}$$

$$\propto |det(\Sigma)|^{-\frac{1}{2}} \cdot p(\Sigma=\tilde{\Sigma}). \tag{2.53}$$

The ratio $\frac{p(B=\tilde{B})}{p(Q=\tilde{Q}|\tilde{\Sigma})}$ is hence a function of Σ which can be hard to characterize analytically, but it is constant along combinations B, Q that imply a given $\tilde{\Sigma}$. In the applications of the paper, this implies that $p(Q=\tilde{Q}|\tilde{\Sigma})$ can be evaluated numerically using

$$p(Q=\tilde{Q}|\tilde{\Sigma}) \propto p(B=\tilde{B}). \tag{2.54}$$

2.A.3.2 The importance sampler

Before discussing the importance sampler, we clarify the notation used so far by noticing that several distributions (for example $p(B|Y)$ and $p(\Sigma|Y)$) have been derived twice: once for the special case associated with the NiWU approach (which requires specifying the hyperparameters $\boldsymbol{\mu}_{NiWU}, V_{NiWU}, d_{NiWU}, S_{NiWU}$), and once for the less restrictive Np(B) (which requires specifying $p(B), \boldsymbol{\mu}_{\pi}, V_{\pi}$). By construction, these sets of functions coincide when the distribution $p(B)$ from the Np(B) approach coincides with the distribution implicit in the independent NiWU approach, equation (2.35), and when $\boldsymbol{\mu}_{NiWU} = \boldsymbol{\mu}_{\pi}$ and $V_{NiWU} = V_{\pi}$. To avoid confusion, the discussion in this section and in the paper uses the following notation:

Np(B) approach:

- $p(B)_{Np(B)}$: prior distribution on B used in the Np(B) approach;
- $p(\Sigma)_{Np(B)}$: prior distribution on Σ associated with the Np(B) approach;
- $p(B|Y)_{Np(B)}$: posterior distribition on B associated with $p(B)_{Np(B)}$ and derived analytically in equation (2.49);
- $p(\Sigma|Y)_{Np(B)}$: posterior distribution on Σ implicit in the Np(B) approach and derived analytically in equation (2.51);
- $p(Q|\Sigma)_{Np(B)}$: distribution on Q implicit in the prior $p(B)_{Np(B)}$ conditioning on the space $\Sigma = BB'$, evaluated as in equation (2.54);

NiWU approach:

- $p(B)_{NiWU}$: prior distribution on B implied by the NiWU approach, derived analytically in equation (2.35) for the independent prior specification;
- $p(\Sigma)_{NiWU}$: prior distribution on Σ explicitly introduced in the NiWU approach, which is the inverse Wishart distribution;
- $p(B|Y)_{NiWU}$: posterior distribition on B implied by the NiWU approach, derived analytically in equation (2.36) for the independent prior specification;
- $p(\Sigma|Y)_{NiWU}$: marginal posterior distribution on Σ implicit in the NiWU approach, derived analytically in equation (2.32) for the independent prior specification;
- $p(Q|\Sigma)_{NiWU}$: distribution on Q explicitly used in the NiWU approach, which is uniform in the space $\mathcal{Q}_\Sigma$ of orthogonal matrices that satisfy the identifying restrictions, given Σ;

Our importance sampling procedure aims to draw from $p(B|Y)_{Np(B)}$ by drawing from $p(\Sigma|Y)_{Np(B)}$ and $p(Q|\Sigma)_{Np(B)}$. To this aim, Stage A of our algorithm uses $p(\Sigma|Y)_{NiWU}$ as an importance function for $p(\Sigma|Y)_{Np(B)}$. While the prior distributions $p(\Sigma)_{Np(B)}$ and $p(\Sigma)_{NiWU}$ may well differ, the corresponding posterior distributions $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$ are closer to each other since Σ is identified and the likelihood function dominates the prior distribution. Accordingly, $p(\Sigma)_{NiWU}$, which can be explored conveniently with a Gibbs sampler, can be used as an importance function for $p(\Sigma|Y)_{Np(B)}$.

Start from a selection of $p(B), \boldsymbol{\mu}_\pi, V_\pi$. Consider the case in which $(\boldsymbol{\mu}_\pi, V_\pi)$ are independent of B, and in which the NiWU approach employed to generate proposal draws is used in its independent prior specification, rather than the conjugate specification. Then select (d_{NiWU} and S_{NiWU}) (see discussion below) and set $\boldsymbol{\mu}_{NiWU} = \boldsymbol{\mu}_\pi$ and $V_{NiWU} = V_\pi$. Following equations (2.51) and (2.32), the weights for Stage A can then be computed as

$$w_d^{\text{stage A}} = \frac{p(\Sigma_d|Y)_{Np(B)}}{p(\Sigma_d|Y)_{NiWU}} \propto \frac{J(B \to \Sigma_d) \cdot \int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB}{|det(\Sigma_d)|^{-\frac{d+k+1}{2}} \cdot e^{-\frac{1}{2} tr[\Sigma_d^{-1} S]}}. \tag{2.55}$$

In equation (2.55), $J(B \to \Sigma_d) = |det(\Sigma_d)|^{-\frac{1}{2}}$ when no zero restriction on B is introduced, and must be computed numerically otherwise. Note that when $(\boldsymbol{\mu}_\pi, V_\pi)$ are independent of B and when proposal draws are obtained from the independent NiWU approach with $\boldsymbol{\mu}_{NiWU} = \boldsymbol{\mu}_\pi$ and $V_{NiWU} = V_\pi$, the term $|det(V_\pi^*)|$ does not contain B and $\boldsymbol{\mu}_{NiWU}^* = \boldsymbol{\mu}_\pi^*$ and $V_{NiWU}^* = V_\pi^*$. Accordingly, several terms in equations (2.51) and (2.32) cancel out in equation (2.55). The algorithm can be modified by generating

proposal draws from the conjugate NiWU specification and/or setting $\boldsymbol{\pi}$ and/or V_π a function of B, but the computation of $w_d^{\text{stage A}}$ must be adjusted accordingly.[18]

To evaluate $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)}dB$ (up to a constant), for each Σ_d we draw enough orthogonal matrices from the algorithm by Rubio-Ramirez et al. (2010) until m_3 are stored that satisfy the restrictions. We then evaluate $p(B)_{Np(B)}$ at $B = h(\Sigma_d)Q_q$ for each $\{Q_q\}_{q=1}^{m_3}$ and each $\{\Sigma_d\}_{d=1}^{m_2}$. Last, we sum along the Q dimension and evaluate $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)}dB$ as the d entry of the obtained $m_2 \times 1$ vector, up to an arbitrary constant scalar. In our application we found that an approximate algorithm that omits the term $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)}dB$ recovers almost the same posterior distribution, see figure 2.42-2.44.

Having obtained draws from $p(\Sigma|Y)_{Np(B)}$ by reweighting draws from $p(\Sigma|Y)_{NiWU}$ using weights $\{w_d^{\text{stage A}}\}$, it remains to draw from $p(Q|\Sigma)_{Np(B)}$ and map draws of Σ and Q into B. Extracting from $p(Q|\Sigma)_{Np(B)}$ is not straightforward. However, the algorithm developed by Rubio-Ramirez et al. (2010) extracts orthogonal matrices uniformly from the full orthogonal space $\mathcal{Q}$, and hence also from the general subset in which the target function $p(Q|\Sigma)_{Np(B)}$ has mass. Hence, the distribution $p(Q|\Sigma)_{NiWU}$ can be used as an importance function for $p(Q|\Sigma)_{Np(B)}$. This requires being able to *evaluate* (rather than draw from) $p(Q|\Sigma)_{Np(B)}$ in correspondence to the proposal extractions, which can be done using $p(Q = \tilde{Q}|\tilde{\Sigma})_{Np(B)} \propto p(B = \tilde{B})$, as discussed above. Last, the uniformity of the importance distribution for Stage B implies $p(Q|\Sigma)_{NiWU} \propto 1$. All in all, it follows that the weights for Stage B can be computed as

$$w_d^{\text{stage B}} = \frac{p(Q|\Sigma)_{Np(B)}}{p(Q|\Sigma)_{NiWU}} \propto p(B)_{Np(B)}. \tag{2.56}$$

In principle, our algorithm can be modified by combining the two separate importance samplers used into a single reweighting procedure. However, since the effective sample size in Stage A and Stage B bears very different interpretations, we find it more

[18] If $\boldsymbol{\pi}$ and/or V_π are a function of B, the equalities $\boldsymbol{\mu}_{NiWU}^* = \boldsymbol{\mu}_\pi^*$ and $V_{NiWU}^* = V_\pi^*$ do not hold and $w_d^{\text{stage A}}$ must be computed as

$$w_d^{\text{stage A}} = \frac{J(B \to \Sigma_d) \cdot |det(V_\pi)|^{-\frac{1}{2}} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{-\boldsymbol{\mu}_\pi^{*\prime} V_\pi^{*-1} \boldsymbol{\mu}_\pi^* + \boldsymbol{\mu}_\pi' V_\pi^{-1} \boldsymbol{\mu}_\pi\right\}} \cdot \int_{\mathcal{B}(\Sigma)} p(B)dB}{|det(\Sigma)|^{-\frac{d_{NiWU}+k+1}{2}} \cdot |det(V_{NiWU}^*)|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{-\boldsymbol{\mu}_{NiWU}^{*\prime} V_{NiWU}^{*-1} \boldsymbol{\mu}_{NiWU}^* + tr\left[\Sigma^{-1} S_{NiWU}\right]\right\}}}.$$

If in addition proposal draws are generated from the conjugate rather than the independent NiWU approach, $w_d^{\text{stage A}}$ should be computed as

$$w_d^{\text{stage A}} = \frac{J(B \to \Sigma_d) \cdot |det(V_\pi)|^{-\frac{1}{2}} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T \otimes \Sigma^{-1}\right)\tilde{\boldsymbol{y}} - \boldsymbol{\mu}_\pi^{*\prime} V_\pi^{*-1} \boldsymbol{\mu}_\pi^* + \boldsymbol{\mu}_\pi' V_\pi^{-1} \boldsymbol{\mu}_\pi\right\}} \cdot \int_{\mathcal{B}(\Sigma)} p(B)dB}{|det(\Sigma)|^{-\frac{d_{NiWU}+k+1}{2}} \cdot \cdot e^{-\frac{1}{2} trace[\Sigma^{-1} S_{NiWU}^*]}},$$

with S_{NiWU}^* computed in accordance with the conjugate NiWU approach.

natural to organize the algorithm into two separate importance samplers. In addition, the computational time is reduced when keeping Stage A and B separate, as a relatively low effective sample size in Stage B requires only running Stage B again with a higher number of proposal draws, without running Stage A again.

Implementing our algorithm requires setting the hyperparameters d_{NiWU}, S_{NiWU}. We propose to set them heuristically to minimize the distance between $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$. We set $d_{NiWU} = k + 2$ and S_{NiWU} such that the expected value of Σ associated with the prior distribution implicit in the importance function equals the prior mean of Σ implicit in $p(B)_{Np(B)}$. This procedure requires to

1. specify $p(B)$;
2. extract m_0 draws from $p(B)$;
3. compute the corresponding draws of $\Sigma = BB'$;
4. set $S_{NiWU} = (d - k - 1) \cdot \frac{\sum_{d=1}^{m_0} \Sigma_d}{m_0}$.

Other parametrizations are also possible.

The simulation exercise in section 2.3 of the paper discusses posterior distributions in the hypothetical case in which the researcher has access to an infinitely large dataset. Our algorithm can be modified to explore such a scenario by running only Stage B using $\Sigma = \Sigma_{true}$ in Step 3, with Σ_{true} the true value of Σ used in the data generating process. The same simulation exercise also explores the distribution of the rotation angle θ implicit in the orthogonal matrices extracted from $p(Q|\Sigma)_{Np(B)}$ and $p(Q|\Sigma)_{NiWU}$. In the bivariate case, since orthogonal matrices can be viewed as Givens rotations

$$Q(\theta) = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}, \tag{2.57}$$

one can explore the distribution on θ implicit in $p(Q|\Sigma)_{Np(B)}$ or $p(Q|\Sigma)_{NiWU}$ by computing $\theta = atan\left(\frac{\bar{Q}_{2,1}}{\bar{Q}_{1,1}}\right)$ for each extraction $\bar{Q}$ from $p(Q|\Sigma)_{Np(B)}$ or $p(Q|\Sigma)_{NiWU}$, with $atan$ the inverse tangent function evaluated at $\frac{\bar{Q}_{2,1}}{\bar{Q}_{1,1}}$, given $\bar{Q}_{i,j}$ the (i,j) entry of $\bar{Q}$. $Q(\theta) = \bar{Q}$ up to sign and ordering of the columns.

Stage A of our importance sampler relies on the convergence of $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$ towards a point mass at Σ_{true} to justify the use of the former distribution as an importance function for the latter. How large the dataset must be to ensure that $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$ are sufficiently close is ultimately an empirical question, and depends not only on the available dataset but also on the prior distributions $p(\Sigma)_{NiWU}$ and $p(B)_{Np(B)}$ used. The simulations on the bivariate VAR model from

section 2.3 find that the ratio of effective draws to proposal draws equals 0.50 (the threshold value used by Herbst and Schorfheide, 2014 to judge whether to change importance function) already for $\tilde{T} = T - p = 52$. In the application to the oil market from section 2.4 of the paper, the dataset effectively has $\tilde{T} = T - p = 405$ observations and delivers a ratio of effective to proposal draws between 0.69 and 0.89 (table 2.14). As a heuristic, back-of-the-envelop calculation, we kept the training sample constant and used the first τ observations from the estimation sample, increasing τ from 85 to $\tilde{T} = 405$. The results, reported in figure 2.49, show that the ratio of effective to proposal draws increases fairly rapidly as the size of the dataset increases. The threshold value of 0.50 is reached for a number of observations of $\tau = 125$ for all three prior specifications.

Table 2.5: Tuning parameters used for our algorithm (Np(B) approach)

			section 2.3	section 2.4
Stage A				
	m_0	draws from $p(B)$ to calibrate proposal	5,000	5,000
	m_1	burn-in draws	50,000	50,000
	m_2	retained draws	50,000	50,000
	m_3	draws of Q to evaluate $\int_{\mathcal{B}(\Sigma)} p(B)_{Np(B)} dB$	30	30
Stage B				
	m_4	maximum number of draws of Q	1,000	1,000
	m_5	proposal draws	80,000	80,000
	Δm_5	incremental for m_5 if $ESS^B < m_5$	40,000	40,000
	m_6	minimum effective sample size accepted	1,000	1,000

2.A.3.3 Diagnostics for the Importance Sampler

Stage B of the importance sampler uses an importance function that features positive mass everywhere in the support of the target function. Accordingly, the favourable condition highlighted by Geweke (1989) in his equation (5) holds, ensuring that the variance of the estimators constructed on the parameters of interest is finite (see also Robert and Casella (2013), chapter 3). However, this condition does not hold for Stage A of the algorithm, which hence requires ensuring that the non-normalised importance weights $\{w_i\}_{i=1}^N$ of size N have a finite variance. As pointed out by Geweke (1989) and Robert and Casella (2013), the finite variance condition is not necessary to ensure consistency of the estimators based on the weighted draws, but it reduces the risk that the sampler performs poorly. In order to assess whether the variance of the weights in Stage A is bounded, we employ a graphical procedure and three diagnostic tests proposed by Koopman et al. (2009).

Graphical Assessment

A measure frequently used to assess the quantitative importance of outliers in importance weights (and hence possible concerns about the finiteness of the variance) is the recursively estimated variance of the weights. Define this estimated variance as $\{v_i\}_{i=1}^N$, where $v_i = Var(w_{1:i})$ is the variance of the first i weights. If outliers do not raise quantitatively relevant concerns, $\{v_i\}_{i=1}^N$ should converge smoothly towards a constant. If, instead, individual outliers dominate the recursive variance, the plot will reveal large jumps. Jumps in $\{v_i\}_{i=1}^N$ are indicative of an unbounded variance of the weights. We use this graphical procedure to assess the performance of the weights in our sampler.

Diagnostic tests: Wald test, Score test and LR test

In addition to using graphical illustrations, we employ the more formal testing procedures by Koopman et al. (2009) to assess the quantitative relevance of outliers. Koopman et al. (2009) propose a Wald test, a score test and a Likelihood Ratio test to test

$$H_0 : \text{the variance is finite} \tag{2.58}$$

against

$$H_1 : \text{the variance is unbounded.} \tag{2.59}$$

To set the stage, note that the problematic part of $\{w_i\}_{i=1}^N$ are the excessively large weights. It is therefore natural to consider only those weights that are larger than a certain threshold u. After specifying u, generate a new random variable

$$Z_i = w_i - u \quad \text{if} \quad w_i > u. \tag{2.60}$$

If $\{w_i\}_{i=1}^N$ are i.i.d. random draws from the same random variable, then, as shown by Pickands et al. (1975), for large N and u the new sequence of random variables $\{Z_i\}_{i=1}^n$ approximates a generalised Pareto distribution with density function

$$f(z|\xi, \beta) = \frac{1}{\beta}\Big(1 + \xi\frac{z}{\beta}\Big)^{-\frac{1}{\xi}-1}. \tag{2.61}$$

The attractive feature of this distribution is that the finiteness of the variance can directly be assessed: If $\xi \leq 0.5$, the variance of Z_i exists, otherwise it is unbounded.

In practice, the threshold u plays a crucial role. We follow Arias et al. (2018) and use five different threshold values, $v_1 = 0.5N$, $v_2 = 0.6N$, $v_3 = 0.7N$, $v_4 = 0.9N$ and $v_5 = 0.99N$ and set $u_j = w_{(v_j)}$, where $w_{(i)}$ are the ordered weights, i.e. $w_1 \leq w_{(2)} \leq$

$... \leq w_{(N)}$. To make the tests operational we use the null and alternative hypotheses

$$H_0 : \xi = 0.5, \tag{2.62}$$

$$H_1 : \xi > 0.5. \tag{2.63}$$

Wald test

The log likelihood function based on the generalised Pareto distribution in (2.61) equals

$$logf(z|\beta,\xi) = -n \cdot log(\beta) - \Big(1 + \frac{1}{\xi}\Big) \sum_{i=1}^{n} log(1 + \xi\beta^{-1} z_i). \tag{2.64}$$

To construct the test statistic, we need to numerically maximise (2.64) with respect to β and ξ to obtain β^{MLE} and ξ^{MLE}, and then construct a test statistic to test for $\xi = 0.5$. In practice, we follow these steps:

1. numerically maximise the unrestricted log likelikood function given in (2.64) to obtain β^{MLE} and ξ^{MLE};

2. construct the test statistic as

$$t = \sqrt{\frac{n}{3\hat{\beta}^{MLE}}}\Big(\hat{\xi}^{MLE} - \frac{1}{2}\Big). \tag{2.65}$$

The statistic t has an approximate standard normal distribution for $n \to \infty$. Large values indicate that ξ exceeds 0.5, which is indicative of unbounded weight variance. Therefore, we reject H_0 of finite variance of the weights if t exceeds the critical value obtained from the standard normal distribution.

Score Test

In order to construct a score test, we need to maximise (2.64) under the null hypothesis of $\xi = 0.5$. We follow these steps:

1. numerically maximise (2.64) with respect to β under the restriction $\xi = 0.5$ to obtain $\beta^{MLE,r}$;

2. since the test statistic is based on the score function of ξ, differentiate (2.64) with respect to ξ and set $\xi = 0.5$ and $\beta = \beta^{MLE,r}$. This gives [19]

$$\hat{s}_r^{\xi} = 4 \sum_{i=1}^{n} log(1 + \frac{z_i}{2\beta^{MLE,r}}) - 6 \sum_{i=1}^{n} \frac{z_i}{2\beta^{MLE,r} + z_i}; \tag{2.66}$$

[19] Please note that there is a typo in Koopman et al. (2009) (page 4) in the corresponding equation.

3. compute

$$s_*^{\xi} = \frac{1}{\sqrt{2n}} \hat{s}_r^{\xi}. \tag{2.67}$$

Since $s_*^{\xi} \to N(0,1)$ for $n \to \infty$, we reject H_0 of finite weight variances if s_*^{ξ} exceeds the critical values obtained from a standard normal distribution.

LR Test

In order to construct a likelihood ratio test statistic, we follow these steps:

1. Estimate the model under the restriction implying infinite weight variance, i.e. maximise (2.64) under the restrictions $\xi \geq 0.5$ to obtain $\tilde{\beta}$ and $\tilde{\xi}$. Compute $log\tilde{f} = f(z|\tilde{\beta}, \tilde{\xi})$;

2. The restricted MLE estimator for β, $\beta^{MLE,r}$, is computed in the previous section. Compute $logf^r = f(z|\beta^{MLE,r}, \xi = 0.5)$;

3. The test statistic is $LR = 2(log\tilde{f} - logf^r)$.

The null hypothesis of finite weight variance is rejected for large values compared to a $\chi(1)$ distribution.

Figure 2.21 and figure 2.50 report the results of the graphical assessment for the applications developed in the paper, while table 2.7 and table 2.16 report the test statistics.

2.A.4 The Dynamic Striated Metropolis-Hastings algorithm by Waggoner et al. (2016)

To assess whether our sampling procedure successfully explores the posterior distributions of the applications in the paper, we sample the same posterior distributions using the sampler developed by Waggoner et al. (2016). We do so because such sampler was intentionally designed to explore posterior distributions that are potentially irregularly shaped and challenging to explore, and hence is appropriate for the application of the paper. While being more computationally demanding, the Dynamic Striated Metropolis-Hastings sampler offers a benchmark against which to compare the extractions of the posterior distribution from our sampler. Section 2.A.4.1 briefly discusses the functioning of the sampler, while section 2.A.4.2 discusses the convergence criteria that we use to assess the performance of the posterior chains.

2.A.4.1 The sampler

The key intuition behind the Dynamic Striated Metropolis-Hastings sampler is that one does not immediately sample $p(B|Y)_{Np(B)}$, which might feature an irregular shape and multiple peaks, but a simpler function. The draws from this starting function are then progressively used to draw from $p(B|Y)_{Np(B)}$. Define $\boldsymbol{\theta}$ the vector including the parameters of interest and the function $f_\lambda(\boldsymbol{\theta})$ as

$$f_\lambda(\boldsymbol{\theta}) = f^s(\boldsymbol{\theta}) \cdot \Big(f^i(\boldsymbol{\theta})\Big)^\lambda \cdot \Big(f^d(\boldsymbol{\theta})\Big)^{1-\lambda}, \tag{2.68}$$

with tempering parameter $\lambda \in [0, 1]$. $f^s(\boldsymbol{\theta})$, $f^i(\boldsymbol{\theta})$ and $f^d(\boldsymbol{\theta})$ are selected such that $f_{\lambda=0}(\boldsymbol{\theta})$ is the kernel of a starting distribution and $f_{\lambda=1}(\boldsymbol{\theta})$ concides with the kernel of the probability distribution that one ultimately wants to explore. As λ increases from 0 to 1, $f_\lambda(\boldsymbol{\theta})$ progressively introduces the elements of $f^i(\boldsymbol{\theta})$ and drops the elements of $f^d(\boldsymbol{\theta})$.

Waggoner et al. (2016) specify $f^s(\boldsymbol{\theta})$, $f^i(\boldsymbol{\theta})$ and $f^d(\boldsymbol{\theta})$ in order to initialize the algorithm from the prior distribution. We depart from Waggoner et al. (2016) and specify $f^s(\boldsymbol{\theta})$, $f^i(\boldsymbol{\theta})$ and $f^d(\boldsymbol{\theta})$ to initialize the algorithm from $p(B|Y)_{NiWU}$. Initializing the algorithm in $p(B|Y)_{NiWU}$ rather than $p(B)_{Np(B)}$ speeds up the posterior sampler, because the starting kernel already introduces information from the data. Then, we progressively convert $p(B|Y)_{NiWU}$ into $p(B|Y)_{Np(B)}$ using the sampler by Waggoner et al. (2016). In our applications, the sampler is required for either four or nine parameters, which is well within the range of parameters in which the Dynamic Striated Metropolis-Hastings sampler performs efficiently.

More precisely, in our applications $\boldsymbol{\theta}$ contains the entries of B, excluding the ones restricted to zero, if any. In sampling $p(B|Y)_{Np(B)}$ from equation (2.8), one could set

$$f^s(\boldsymbol{\theta}) = p(B)_{Np(B)}, \tag{2.69}$$

$$f^i(\boldsymbol{\theta}) = |det(B)|^{-T} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T \otimes (BB')^{-1}\right)\tilde{\boldsymbol{y}} - \boldsymbol{\mu}_\pi^{*\prime} {V_\pi^*}^{-1} \boldsymbol{\mu}_\pi^*\right\}}, \tag{2.70}$$

$$f^d(\boldsymbol{\theta}) = 1. \tag{2.71}$$

This specification follows Waggoner et al. (2016) in initializing the algorithm at the prior distribution. We propose, instead, to start from a function that allows exploiting the computational advantage of the NiWU approach from section 2.2.2 of the paper. Since the posterior distribution corresponding to the NiWU approach is available analytically, equation (2.32), we use this function as the starting point of the algorithm

and set

$$f^s(\boldsymbol{\theta}) = |det(B)|^{-T} \cdot |det(V_\pi^*)|^{\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{\tilde{\boldsymbol{y}}'\left(I_T \otimes (BB')^{-1}\right)\tilde{\boldsymbol{y}} - \boldsymbol{\mu}_\pi^{*\prime} V_\pi^{*-1} \boldsymbol{\mu}_\pi^*\right\}}, \tag{2.72}$$

$$f^i(\boldsymbol{\theta}) = p(B)_{Np(B)} \cdot |det(V_\pi)|^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}\left\{\boldsymbol{\mu}_\pi' V_\pi^{-1} \boldsymbol{\mu}_\pi\right\}}, \tag{2.73}$$

$$f^d(\boldsymbol{\theta}) = |det(B)|^{-(d+k)} e^{-\frac{1}{2}\left\{vec(B^{-1})'(S \otimes I_k) vec(B^{-1})\right\}}. \tag{2.74}$$

We set d and S as in our algorithm, see section 2.A.3.2. If zero restrictions on B are introduced, either equation (2.74) is modified to numerically account for the volume element featuring also zero restrictions or the algorithm is initialized in $p(B)_{Np(B)}$.

Once the functions $f^s(\boldsymbol{\theta})$, $f^i(\boldsymbol{\theta})$ and $f^d(\boldsymbol{\theta})$ are selected, define the target function

$$log[f_{\lambda_h}(\boldsymbol{\theta})] = log[f^s(\boldsymbol{\theta})] + \lambda_h \cdot log[f^i(\boldsymbol{\theta})] + (1 - \lambda_h) \cdot log[f^d(\boldsymbol{\theta})], \tag{2.75}$$

as the logarithm of the tempered function $f_{\lambda_h}(\boldsymbol{\theta})$ at stage h, with $h = 1, ..., H$ and H the total number of stages. To the extent that, among other requirements, a sufficient number of stages is used, the target function at stage $h-1$ is sufficiently close to the target function at stage h. This makes the draws representative of the target function at stage $h-1$ a useful point of departure to numerically explore the target distribution at stage h.

Within this sequential approach, Waggoner et al. (2016) propose sampling $log[f_{\lambda_h}(\boldsymbol{\theta})]$ using a modified Metropolis-Hastings algorithm.[20] In the general stage h, the algorithm can be summarized in the following steps:

1. start stage h with $N \cdot G$ draws $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$, which are representative of the target distribution at stage $h-1$. If $h = 1$, then $\{\boldsymbol{\theta}_d^{(0)}\}_{d=1}^{N \cdot G}$ are drawn from the starting function, otherwise, they are computed at the end of stage $h-1$;

2. evaluate the function $log[f^i(\boldsymbol{\theta})]$ at each $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$ and group draws $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$ into M 'striations' (subsets), depending on the corresponding value of $log[f^i(\boldsymbol{\theta})]$;

3. for each $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$ compute weights $\tilde{\omega}_d = \frac{f_{\lambda_h}(\boldsymbol{\theta}_d^{(h-1)})}{f_{\lambda_{h-1}}(\boldsymbol{\theta}_d^{(h-1)})}$. As with importance sampling techniques, $\omega_d = \frac{\tilde{\omega}_d}{\sum_{d=1}^{N \cdot G} \tilde{\omega}_d}$, $d = 1, .., N \cdot G$ allow reweighting the draws from the previous stage such that they become representative of the target distribution of the current stage, provided that the effective sample size does not shrink excessively;

[20] See, for example, Herbst and Schorfheide (2014) for an alternative approach to sequential Monte Carlo samplers.

4. use $\{\boldsymbol{\theta}_d^{(h-1)}, \omega_d\}_{d=1}^{N \cdot G}$ to compute numerically the variance Ω_h of the target function at stage h;

5. explore the target function $log[f_{\lambda_h}(\boldsymbol{\theta})]$ as follows. For each group g, set the initial draw $\boldsymbol{\theta}_{old}$ to a random draw from $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$, extracted with replacement using $\{\omega_d\}_{d=1}^{N \cdot G}$. Then, with probability p, set $\boldsymbol{\theta}_{new} = \boldsymbol{\theta}_{old} + \boldsymbol{\theta}_{shock}$ with $\boldsymbol{\theta}_{shock}$ a multivariate zero-mean normal random variable with variance $c_h \cdot \Omega_h$, while with probability $1 - p$ set $\boldsymbol{\theta}_{new}$ to a randomly extracted draw from the subset of $\{\boldsymbol{\theta}_d^{(h-1)}\}_{d=1}^{N \cdot G}$ from the striation associated with function $log[f^i(\boldsymbol{\theta})]$ evaluated at $\boldsymbol{\theta}_{old}$. Accept $\boldsymbol{\theta}_{new}$ with probability $\min\left\{1, \frac{f_{\lambda_h}(\boldsymbol{\theta}_{new})}{f_{\lambda_h}(\boldsymbol{\theta}_{old})}\right\}$ if $\boldsymbol{\theta}_{new}$ was generated from the random walk extraction, and with probability $\min\left\{1, \frac{f_{\lambda_h}(\boldsymbol{\theta}_{new})}{f_{\lambda_{h-1}}(\boldsymbol{\theta}_{old})} \frac{f_{\lambda_h}(\boldsymbol{\theta}_{old})}{f_{\lambda_{h-1}}(\boldsymbol{\theta}_{new})}\right\}$ if $\boldsymbol{\theta}_{new}$ was randomly selected from the strations at the previous stage. Continue for $N \cdot \tau$ iterations;

6. store one every τ draws for each group and collect the $N \cdot G$ draws $\{\boldsymbol{\theta}_d^{(h)}\}_{d=1}^{N \cdot G}$. Use $\{\boldsymbol{\theta}_d^{(h)}\}_{d=1}^{N \cdot G}$ to initialize the next stage $h + 1$. If $h = H$, the last stage has been reached, and $\{\boldsymbol{\theta}_d^{(h)}\}_{d=1}^{N \cdot G}$ are interpreted as posterior draws from the desired distribution.

To make the above algorithm operational, we need to set several tuning parameters. Following Waggoner et al. (2016), we set $p = 1/(10\tau)$. We then set the parameter c_h at each stage following the guidance discussed by Waggoner et al. (2016) in Appendix A, hence ensuring, within each stage, an acceptance ratio between 0.20 and 0.30 from a preliminary Metropolis-Hastings algorithm with K iterations. We set the progression of the tempering parameter as in Herbst and Schorfheide (2014), using $\lambda_h = \left(\frac{h-1}{H-1}\right)^2$. It remains to calibrate the number of stages H, the number of groups G, the effective number of iterations N within each group, the number K affecting the number of iterations $K \cdot G$ for the calibration of the parameter c_h, the frequency τ at which draws stored, and the number of striations M. These parameters affect the time required for the algorithm to run. For the applications in this paper, we speed up posterior sampling by setting these parameters to relatively low values, under the following conditions:

1. that the effective sample size computed for the importance sampling from one stage to the next one never falls below 80%. If this occurs a higher number of stages H is used;

2. that the convergence criteria discussed in the next section of this appendix confirm that the draws at stage h have converged to the distribution of the target function at stage h. If this is not the case, a higher number of effective draws N and/or groups G is used.

Table 2.6 reports the values of the tuning parameters used in the applications.

Table 2.6: Tuning parameters used for the DSMH algorithm

		section 2.3	section 2.4
H	number of stages	6	8
G	number of groups	8	8
N	number of MH iterations per group	8,000	12,000
K	number of iterations to calibrate c	1,000	1,000
τ	frequency of storage of iterations	1	1
M	number of striations	10	10

2.A.4.2 Convergence criteria

Consider the chain $\{\boldsymbol{\theta}_d\}_{d=1}^{N \cdot G}$, with $\boldsymbol{\theta}_d = (\theta_{1,d}, .., \theta_{j,d}, .., \theta_{\kappa,d})'$ of dimension $\kappa \times 1$. We employ four convergence criteria in order to assess if the chain has converged in distribution, namely the converge criteria developed by Geweke (1992), by Raftery and Lewis (1992), by Gelman and Rubin (1992) and by Brooks and Gelman (1998). Intuitively, the criteria used operate by assessing whether the series has excessive auto-dependence (which indicates that the draws are not from a stationary distribution) and whether it depends on the starting point (which indicates that the chain is not long enough).

The convergence criteria that we use can be classified according to two main features. First, whether the convergence of each parameter is assessed in isolation from the remaining $\kappa - 1$ parameters or jointly (i.e. whether the object of interest is $\theta_{j,d}$ or $\boldsymbol{\theta}_d$). Second, whether the series of $N \cdot G$ draws are considered in a long chain from a single starting value or in multiple chains from multiple starting points (i.e. whether the object of interest is $\{\theta_{j,d}\}_{d=1}^{N \cdot G}$ or $\{\theta_{j,d}\}_{d=1}^{N}$, $\{\theta_{j,d}\}_{d=N+1}^{2N}$, ..., $\{\theta_{j,d}\}_{d=(G-1)N+1}^{N \cdot G}$ for univariate chains, and $\{\boldsymbol{\theta}_d\}_{d=1}^{N \cdot G}$ or $\{\boldsymbol{\theta}_d\}_{d=1}^{N}$, $\{\boldsymbol{\theta}_d\}_{d=N+1}^{2N}$, ..., $\{\boldsymbol{\theta}_d\}_{d=(G-1)N+1}^{N \cdot G}$ for multivariate chains). The statistic by Brooks and Gelman (1998) considers the multidimensional objects, while the remaining criteria consider univariate objects. The criteria by Geweke (1992) and Raftery and Lewis (1992) consider single chains, while the criteria by Gelman and Rubin (1992) and Brooks and Gelman (1998) consider multiple chains. For a detailed comparative review of convergence criteria for Markov Chain Monte Carlo mechanisms see Cowles and Carlin (1996) and Brooks and Roberts (1998).

Geweke (1992)

The univariate approach by Geweke (1992) assesses the convergence of each parameter of the series in isolation, using the series $\{\theta_{j,d}\}_{d=1}^{N \cdot G}$ for each parameter j. The assessment is based on a comparison of means across different parts of the chain. If the means are close to each other, the procedure detects convergence.

To run the test we proceed in four steps:

1. Extract the first 10% and the last 40% of the draws of $\{\theta_{j,d}\}_{d=1}^{N\cdot G}$, i.e. $\{\theta_{j,d}\}_{d=1}^{0.10\cdot N\cdot G}$ and $\{\theta_{j,d}\}_{d=0.60\cdot N\cdot G}^{N\cdot G}$;

2. For each subseries, compute the mean and the standard deviation and call them $\hat{\mu}_{first}$, $\hat{\mu}_{last}$, $\hat{\sigma}_{first}$ and $\hat{\sigma}_{last}$;

3. Compute the test statistic

$$CD = \frac{\hat{\mu}_{first} - \hat{\mu}_{last}}{\frac{\hat{\sigma}_{first}}{\sqrt{0.1NG}} + \frac{\hat{\sigma}_{last}}{\sqrt{0.4NG}}}. \tag{2.76}$$

   Under the conditions mentioned in Geweke (1992), CD has an asymptotic standard normal distribution;

4. Compute the p-value.

The final statistic of the test is the p-value associated with the statistic CD. A p-value below the significant level indicates that the null hypothesis of convergence, captured by the equality of means across the chain, can be rejected, and hence that the series has not converged.

Raftery and Lewis (1992)

The approach by Raftery and Lewis (1992), like the one by Geweke (1992), investigates one long univariate chain of draws for one parameter in isolation, $\{\theta_{j,d}\}_{d=1}^{N\cdot G}$. The main objects of interest are the quantiles of the probability distribution for the parameter j. The method assesses if the chain is long enough to get precise estimates of quantiles of this distribution.

To define the notion of closeness, three values have to be specified by the user: s, q and r. If the interest lies in $q_{j,0.025}$, the 0.025 quantile of the posterior of a parameter θ_j, then $q = 0.025$. If one exerts 95% of the posterior draws to lie in an interval of +/- 0.0125 around the true 0.025 quantile, then $s = 0.95$ and $r = 0.0125$. These specifications are standard for output from an MCMC chain. The implementation of the algorithm for each parameter j proceeds in 4 steps:

1. Transform $\{\theta_{j,d}\}_{d=1}^{N\cdot G}$ into a dichotomous random variable Z_d:

$$Z_d = \begin{cases} 1 & \text{if } \theta_{j,d} < q_{0.025}, \\ 0 & \text{otherwise;} \end{cases} \tag{2.77}$$

2. Write the matrix of transition probabilities for Z_d conditioning on the previous state,

$$\mathcal{P} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix},$$

with $\alpha = P(Z_{d+1} = 1|Z_d = 0)$ and $\beta = P(Z_{d+1} = 0|Z_d = 1)$. The unconditional probabilities of being in one state or another are

$$\pi_0 = P(\theta_{j,d} < q_{0.025}) = P(Z_d = 0) = \frac{\beta}{\alpha + \beta} \tag{2.78}$$

$$\pi_1 = P(\theta_{j,d} \geq q_{0.025}) = P(Z_d = 1) = 1 - \pi_0 = \frac{\alpha}{\alpha + \beta} \tag{2.79}$$

3. Approximate the probability that a draw of the parameter is smaller than the quantile of interest as

$$P(\theta_{j,d} < q_{0.025}) \approx \bar{Z}_{NG,j} = \frac{1}{NG} \sum_{d=1}^{NG} Z_d. \tag{2.80}$$

As shown by Raftery and Lewis (1992), $\bar{Z}_{NG}$ is approximately normally distributed with mean $q_{0.025}$ and variance $\frac{1}{NG}\frac{(2-\alpha-\beta)\alpha\beta}{(\alpha+\beta)^3}$;

4. Compute the optimal length of the chain as the length that ensures $P(q - r \leq \bar{Z}_{NG} \leq q + r)$ using

$$n^* = \frac{(2-\alpha-\beta)\alpha\beta}{(\alpha+\beta)^3} \left\{ \frac{\Phi^{-1}(\frac{1}{2}(s+1))}{r} \right\}^2 \tag{2.81}$$

The key statistic of the test is n^*, which has an intuitive interpretation: it is the minimum number of draws we need for the desired level of accuracy of the quantile q (given by r and s). If $N \cdot G$ is lower than n^*, this suggests that the chain length needs to be increased.

Gelman and Rubin (1992)

The convergence diagnostic by Gelman and Rubin (1992) uses multiple univariate chains, $\{\theta_{j,d}\}_{d=1}^{N}, \{\theta_{j,d}\}_{d=N+1}^{2N}, ..., \{\theta_{j,d}\}_{d=(G-1)N+1}^{N \cdot G}$. If the chains have converged, then they should not depend on starting values any more. The convergence statistic is based on a comparison of between-sequence variance and within sequence variance. The procedure consists of four steps:

1. Compute the variance of the mean of each sequence ("between-sequence variance") as

$$B = \frac{1}{G-1}\sum_{g=1}^{G}(\bar{\theta}_{j,\cdot,g} - \bar{\theta}_j)^2, \tag{2.82}$$

where $\bar{\theta}_{j,\cdot,g}$ is the mean of $\theta_{j,d}$ within the g-th chain and $\bar{\theta}_j$ is the mean across all chains for parameter j.

2. Compute the mean across sequences of the variances within sequence (the "within-sequence variance") as

$$W = \frac{1}{G(N-1)}\sum_{g=1}^{G}\sum_{n=1}^{N}(\theta_{j,n,g} - \bar{\theta}_{j,\cdot,g})^2; \tag{2.83}$$

3. Estimate the overall variance as

$$\hat{V} = \hat{\sigma}_+^2 + \frac{B}{G}, \tag{2.84}$$

with

$$\hat{\sigma}_+^2 = \frac{N-1}{N}W + B; \tag{2.85}$$

4. Compute the statistic

$$\hat{R} = \frac{\hat{V}}{W}. \tag{2.86}$$

The key statistic of the test is $\hat{R}$. As a rule-of-thumb, $\hat{R}$ should be below 1.2 to assert that the chain has converged.

Brooks and Gelman (1998)

The statistic by Brooks and Gelman (1998) is a multivariate extension of Gelman and Rubin (1992) and requires different chains of a multivariate series, $\{\boldsymbol{\theta}_d\}_{d=1}^{N}, \{\boldsymbol{\theta}_d\}_{d=N+1}^{2N}$, ..., $\{\boldsymbol{\theta}_d\}_{d=(G-1)N+1}^{N\cdot G}$. Intuitively, as in the test by Gelman and Rubin (1992), the approach by Brooks and Gelman (1998) builds the analysis by comparing the between-chain and within-chain variances. The test builds on the multivariate extension of the steps used for the the approach by Gelman and Rubin (1992):

1. Compute the variance of the mean of each sequence ("between-sequence variance") as

$$B = \frac{1}{G-1}\sum_{g=1}^{G}(\bar{\boldsymbol{\theta}}_g - \bar{\boldsymbol{\theta}})(\bar{\boldsymbol{\theta}}_g - \bar{\boldsymbol{\theta}})'; \tag{2.87}$$

2. Compute the mean across sequences of the variances within sequence (the "within-sequence variance") as

$$W = \frac{1}{G(N-1)} \sum_{g=1}^{G} \sum_{n=1}^{N} (\boldsymbol{\theta}_{n,g} - \bar{\boldsymbol{\theta}}_g)(\boldsymbol{\theta}_{n,g} - \bar{\boldsymbol{\theta}}_g)'; \tag{2.88}$$

3. Estimate the overall variance as

$$\hat{V} = \frac{N-1}{N} W + \Big(1 + \frac{1}{G}\Big) B; \tag{2.89}$$

4. Compute the (scalar) distance measure between these two matrices as

$$\hat{R}^{mult} = \frac{N}{N-1} + \frac{G+1}{G} \lambda_1, \tag{2.90}$$

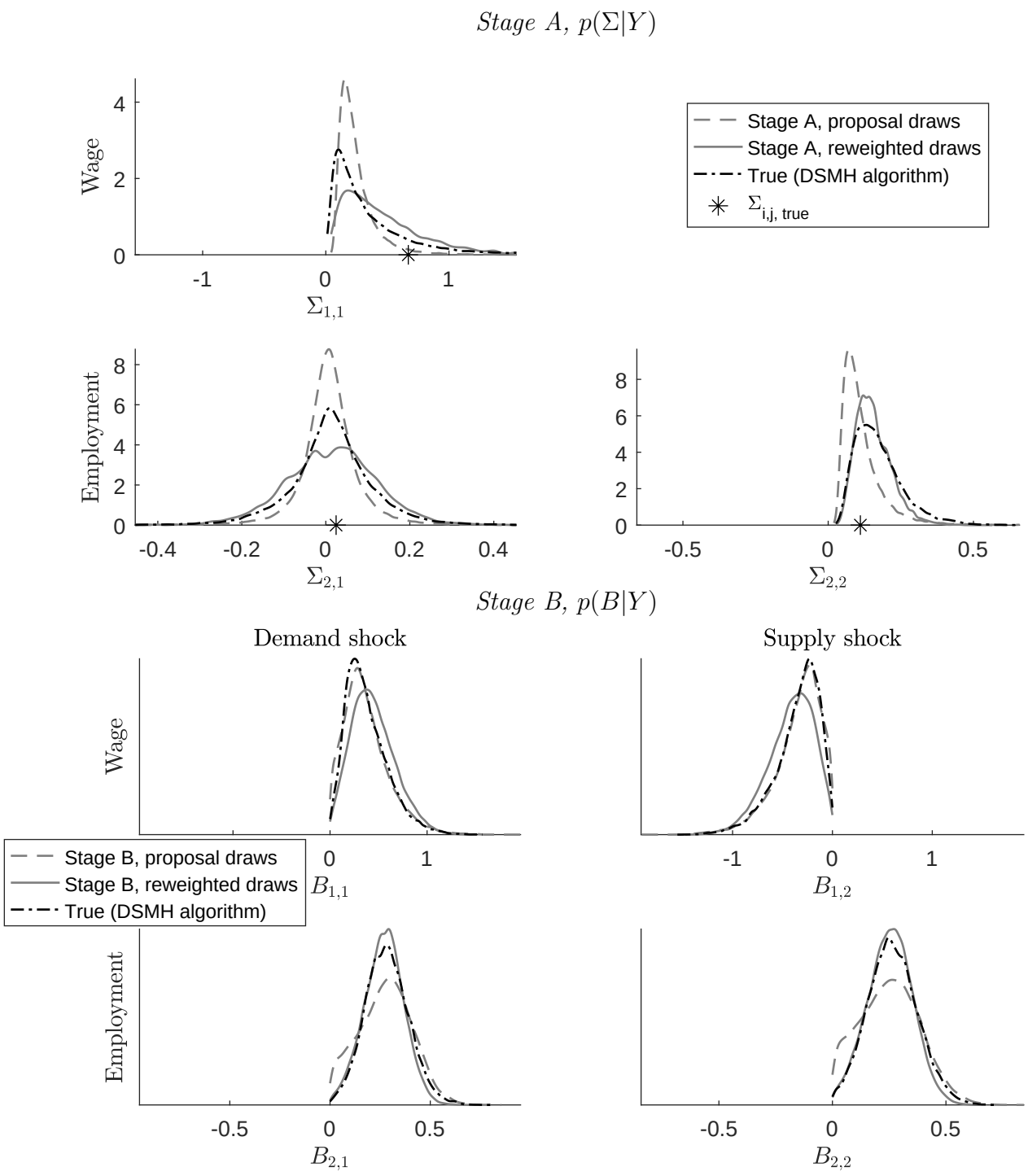
where λ_1 is the largest eigenvalue of the matrix $W^{-1}B$.

The final statistic of the test is $\hat{R}^{mult}$. As for the approach by Gelman and Rubin (1992), the rule-of-thumb prescribes that $\hat{R}^{mult}$ is below 1.2 in order to assert convergence.

Table 2.8 to table 2.12 and table 2.17 to table 2.19 report the convergence diagnostics for the two applications developed in the paper.

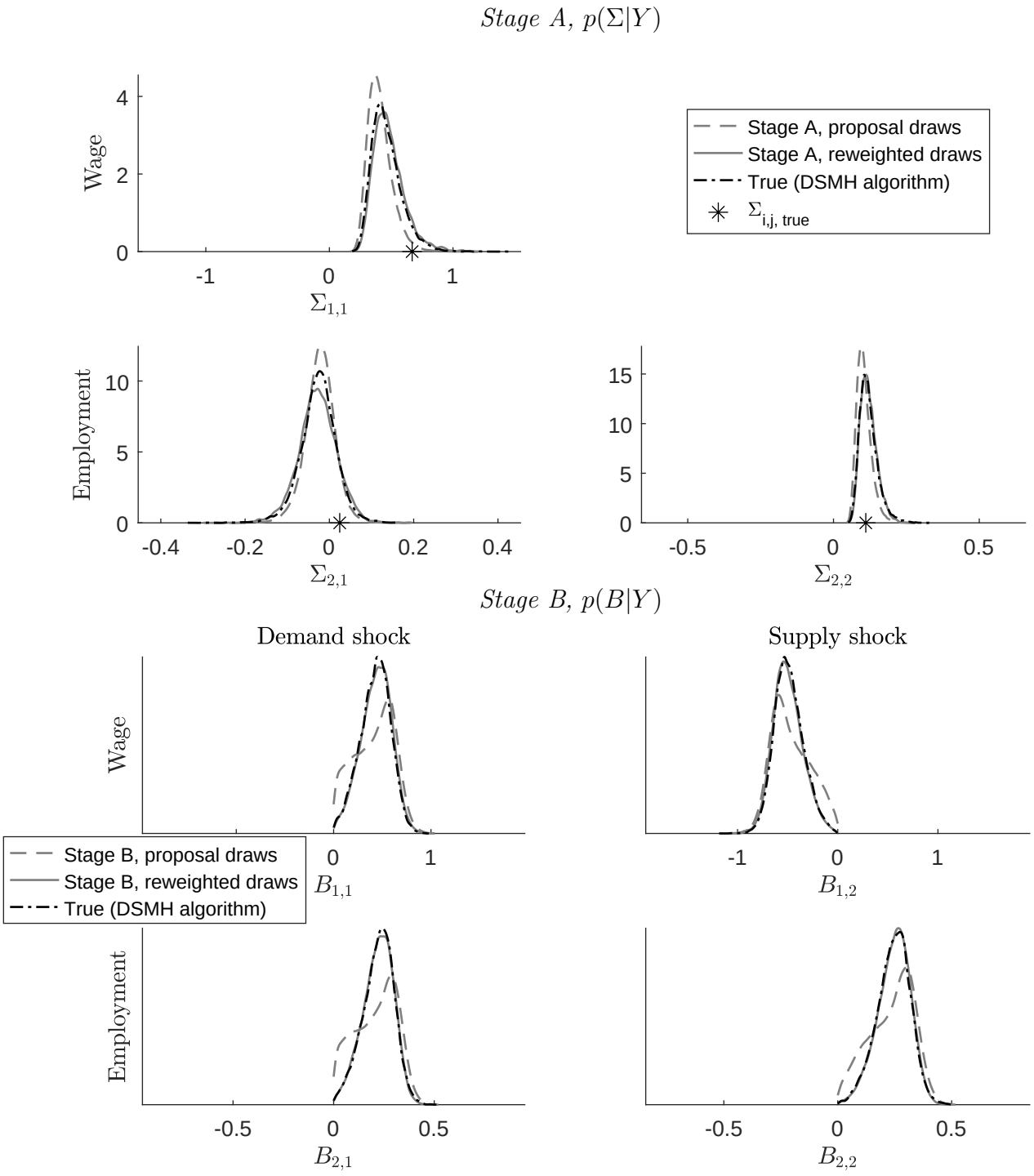
2.A.5 Additional tables/figures for section 2.3

Figure 2.7: Illustration of our algorithm, $T = 30$



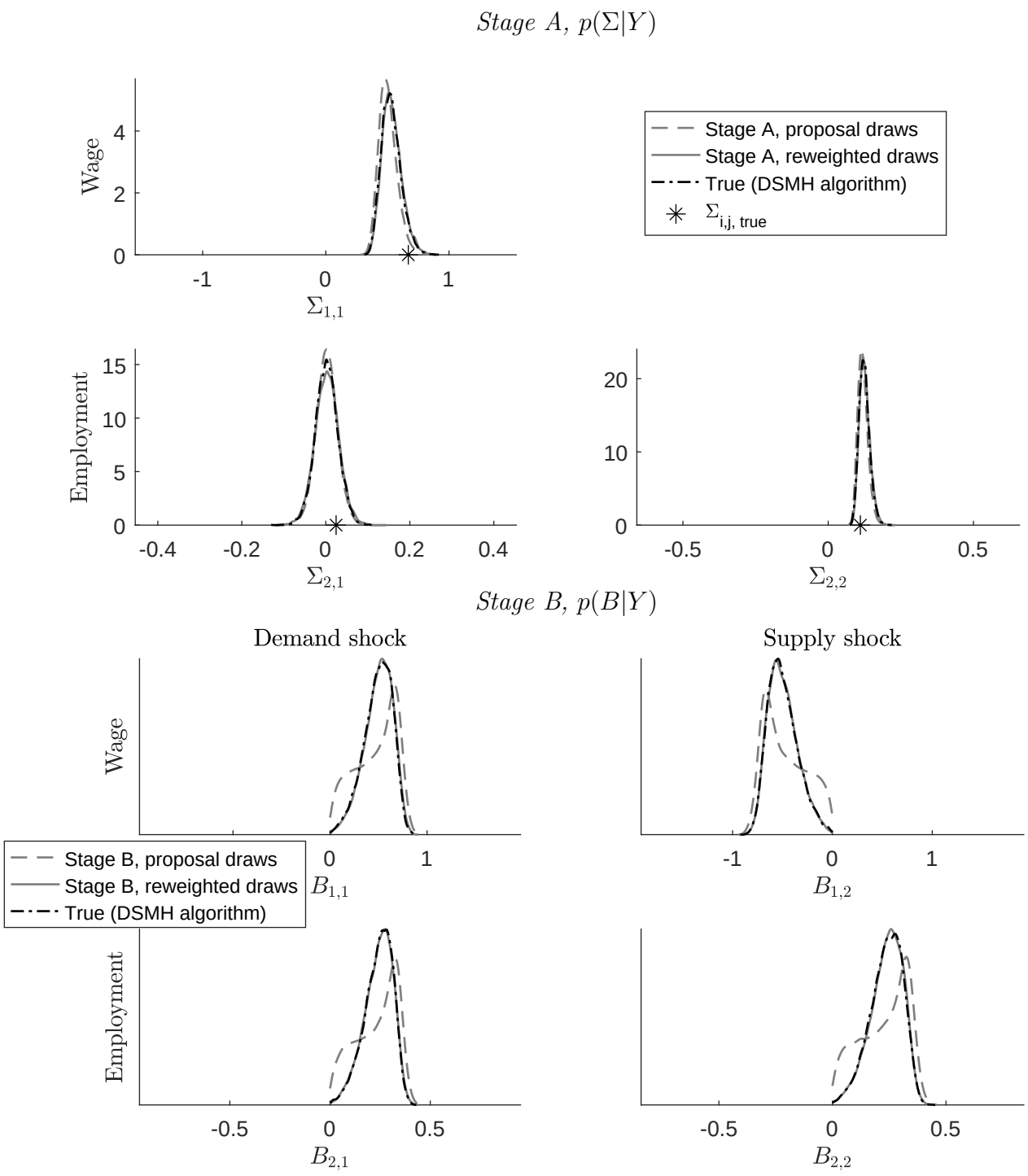
Note: In the upper panel, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the lower panel, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$.

Figure 2.8: Illustration of our algorithm, $T = 60$



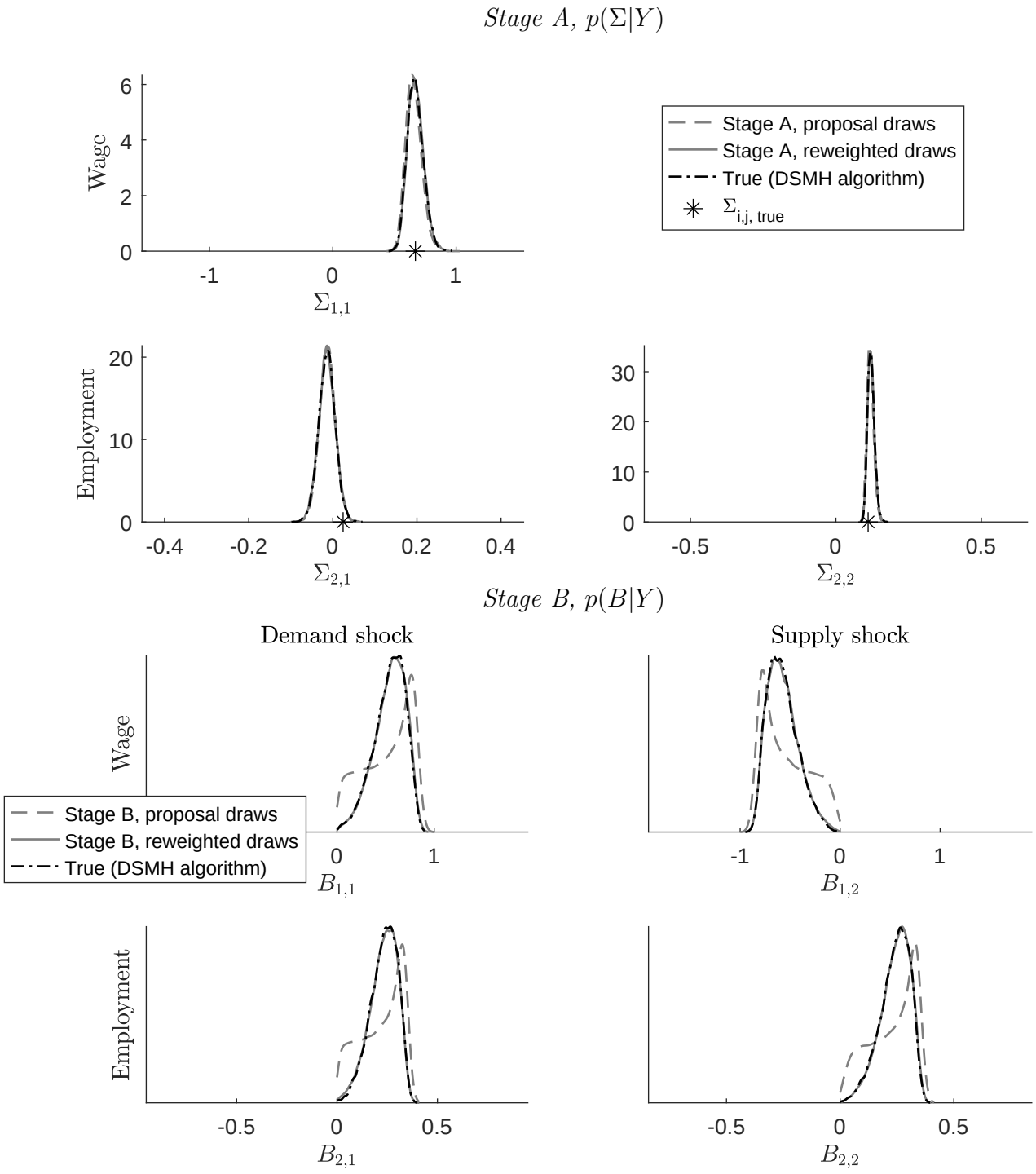
Note: In the upper panel, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the lower panel, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$.

Figure 2.9: Illustration of our algorithm, $T = 120$



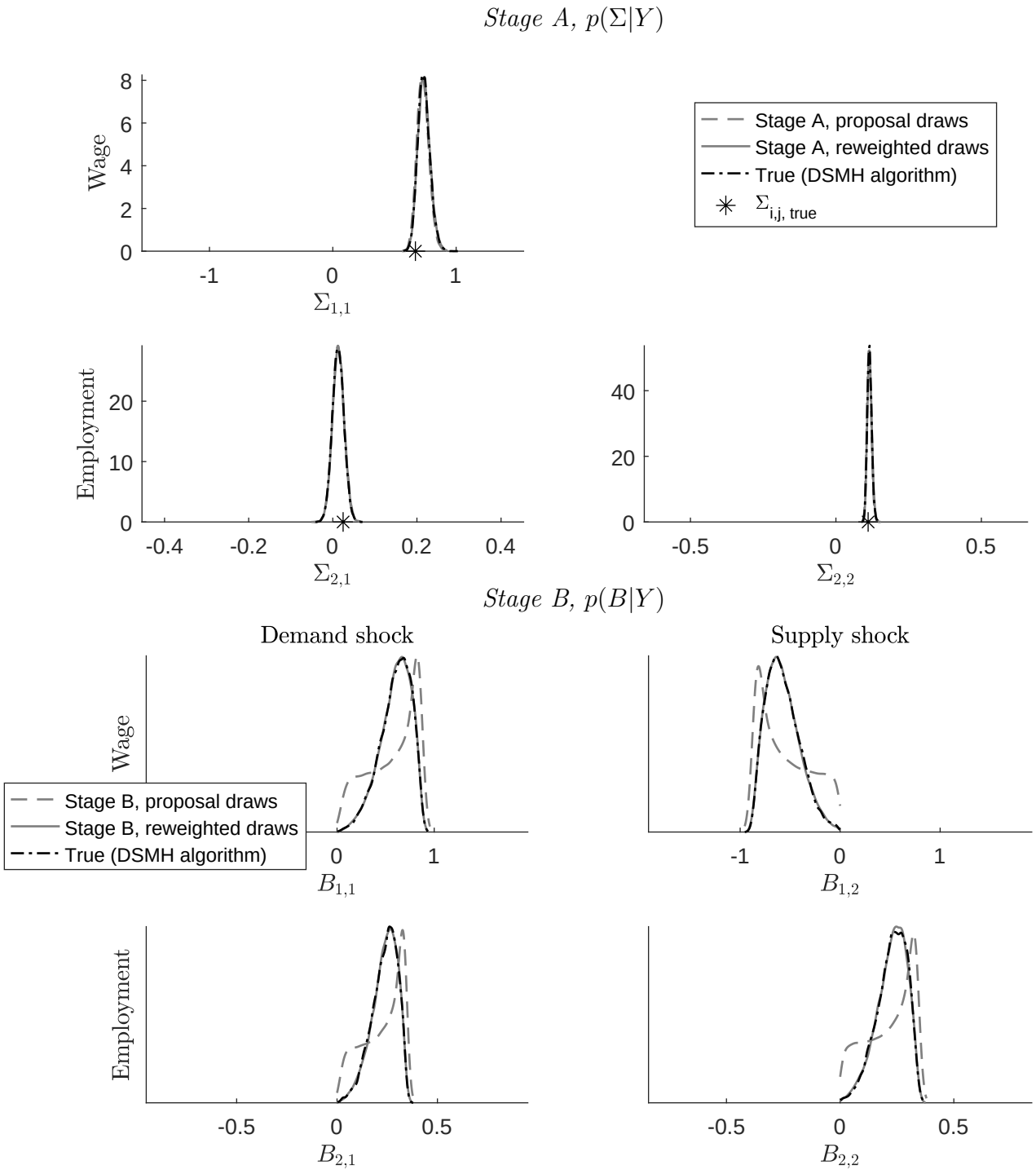
Note: In the upper panel, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the lower panel, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$.

Figure 2.10: Illustration of our algorithm, $T = 240$



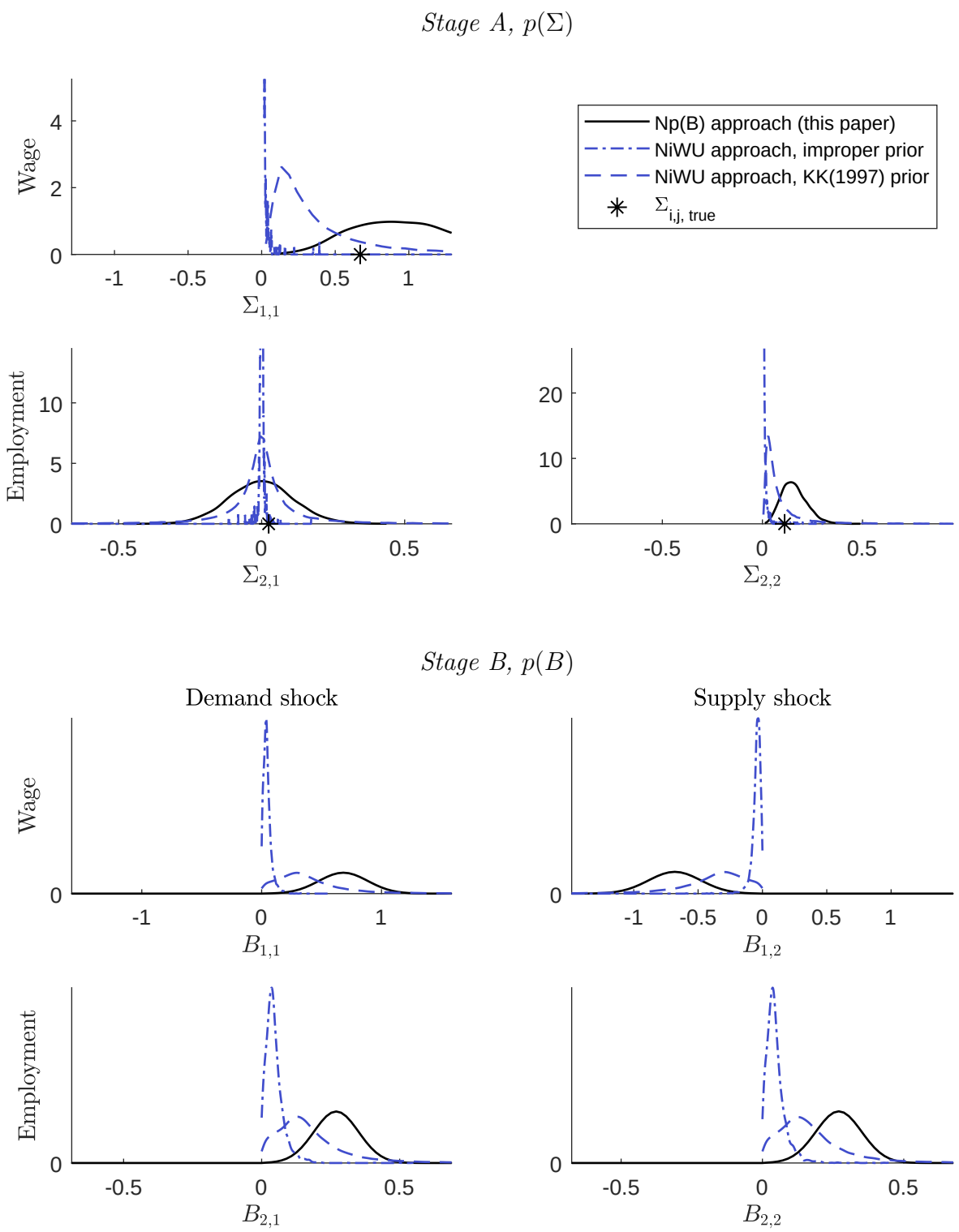
Note: In the upper panel, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the lower panel, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$.

Figure 2.11: Illustration of our algorithm, $T = 480$



Note: In the upper panel, the proposal draws are obtained from Step 1 in our algorithm, the reweighted draws correspond to the same draws reweighted using weights from Step 2, while the draws associated with the Dynamic Striated Metropolis-Hastings algorithm are obtained indirectly after running such algorithm on $p(B|Y)_{Np(B)}$. In the lower panel, the proposal draws correspond to draws obtained from Step 6 of our algorithm, the reweighted draws are obtained from Step 8, and the remaining draws are associated with the Dynamic Striated Metropolis-Hastings algorithm run on $p(B|Y)_{Np(B)}$.

Figure 2.12: Comparing our approach to the NiWU approach: priors



Note: Since the prior distribution for the NiWU approach with $d = 0, S = 0$ is an improper function, in this figure we approximate it by setting $d = k + 2, S = 0.01 I_k$. See also figure 2.13.

Figure 2.13: More on the approximate improper prior in the NiWU approach

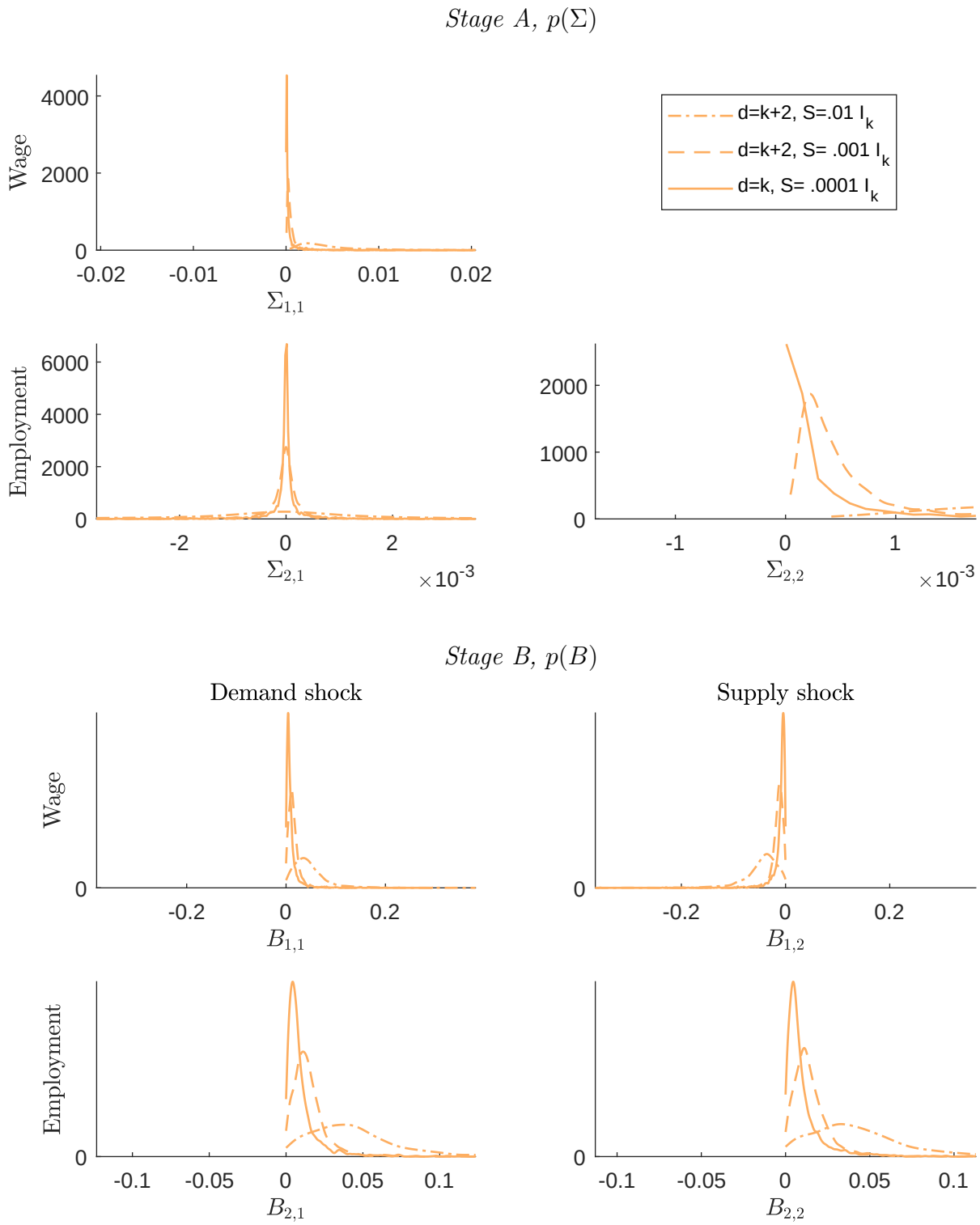


Figure 2.14: Comparing our approach to the NiWU approach, posteriors for $T = 30$

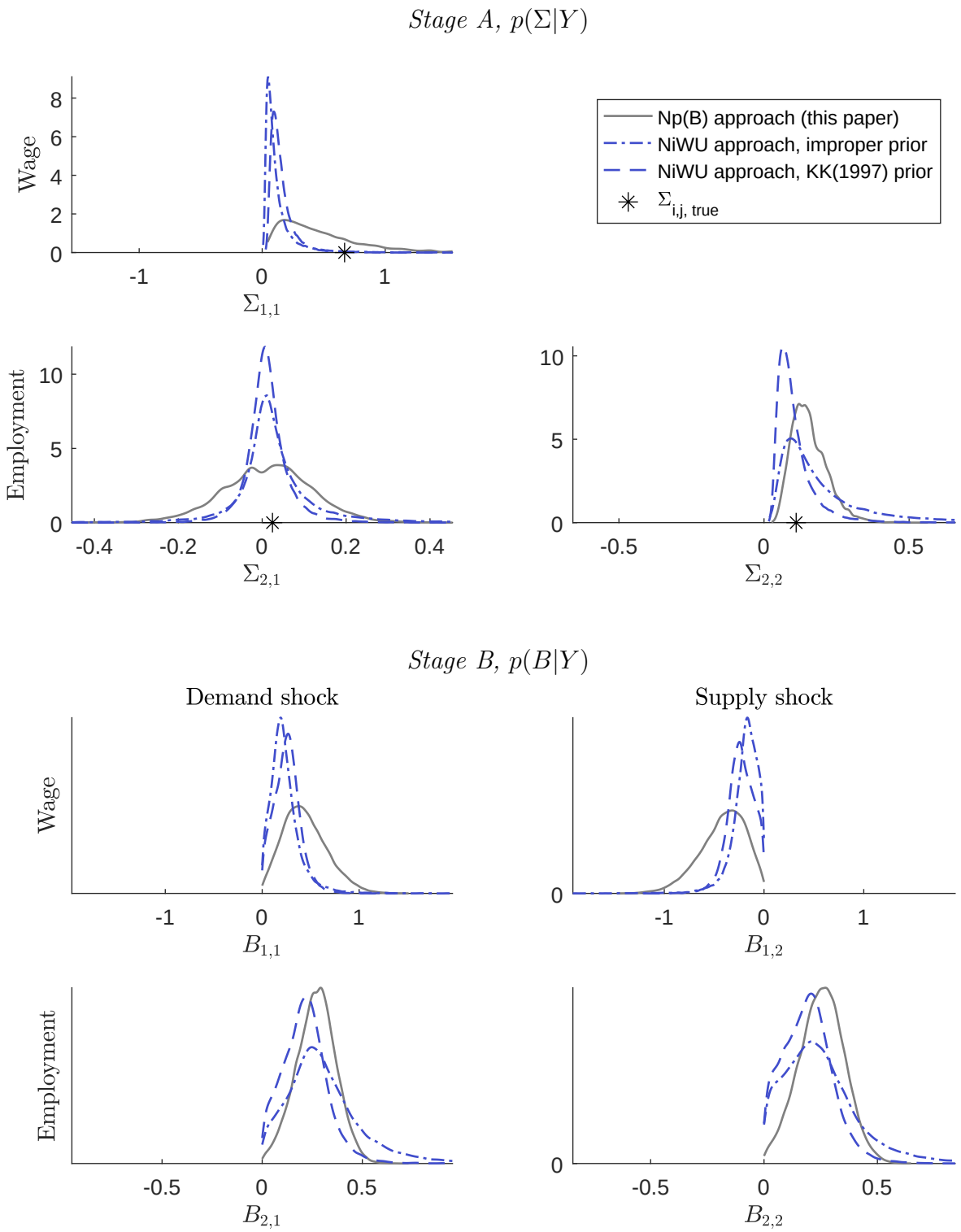


Figure 2.15: Comparing our approach to the NiWU approach, posteriors for $T = 60$

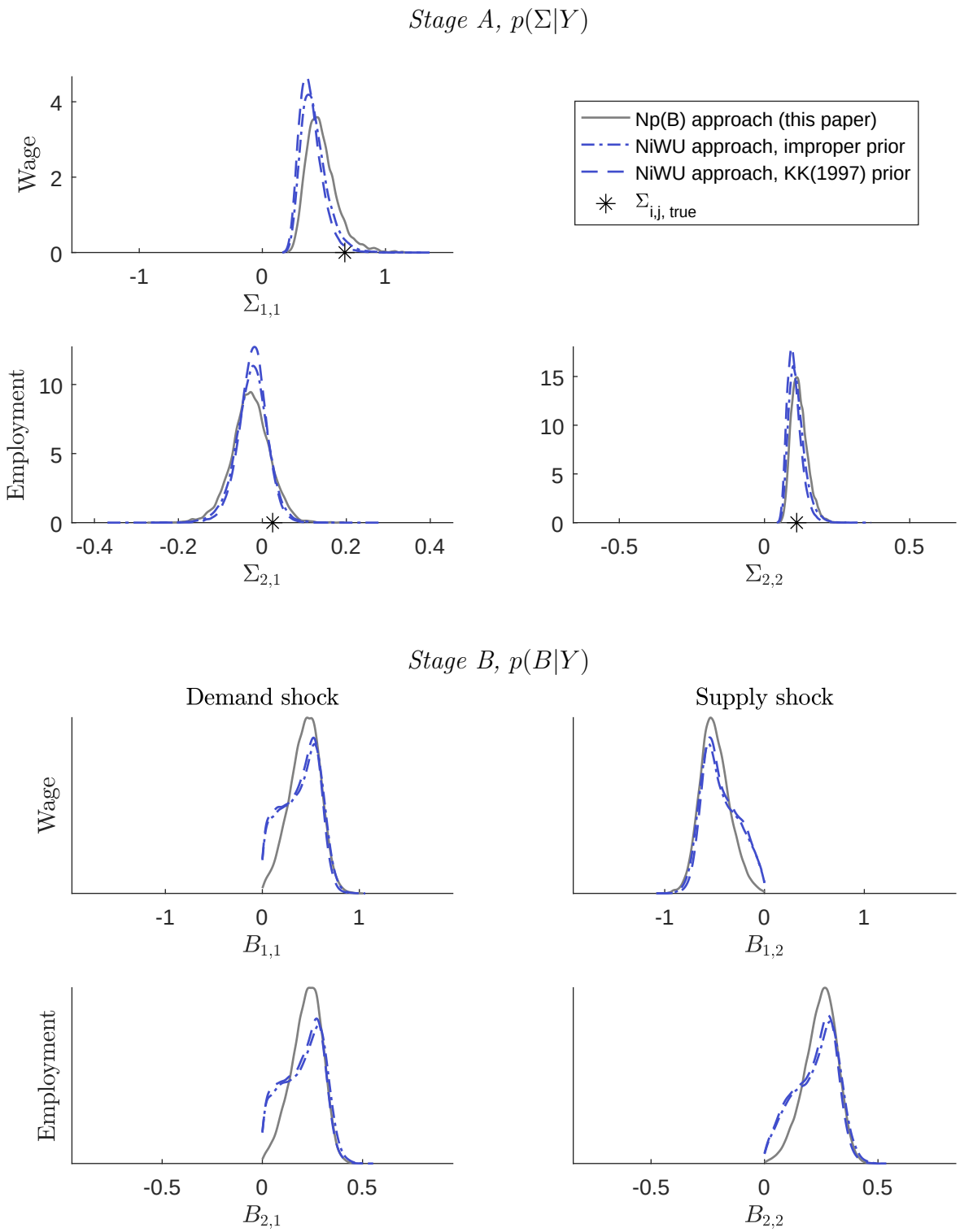


Figure 2.16: Comparing our approach to the NiWU approach, posteriors for $T = 120$

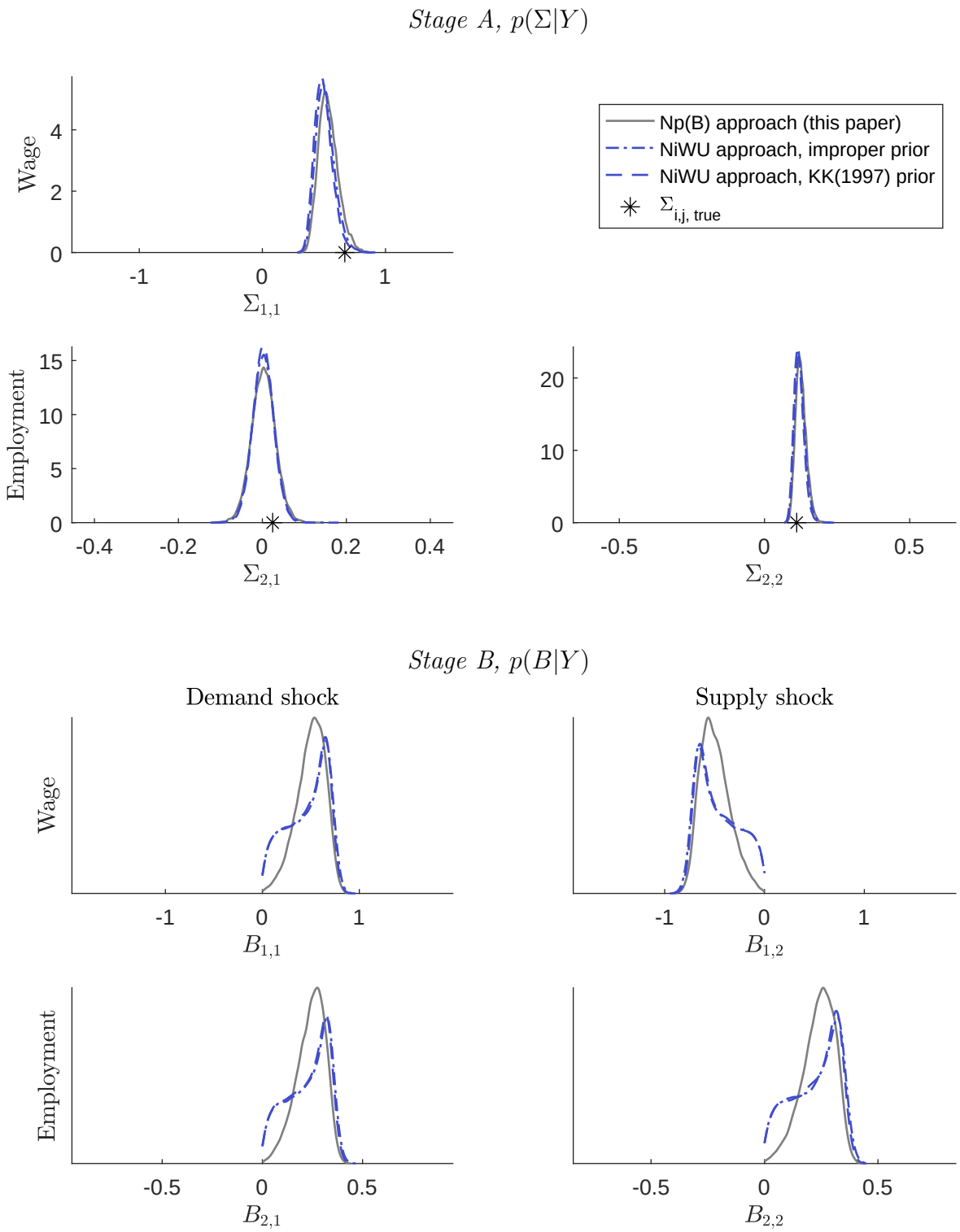


Figure 2.17: Comparing our approach to the NiWU approach, posteriors for $T = 240$

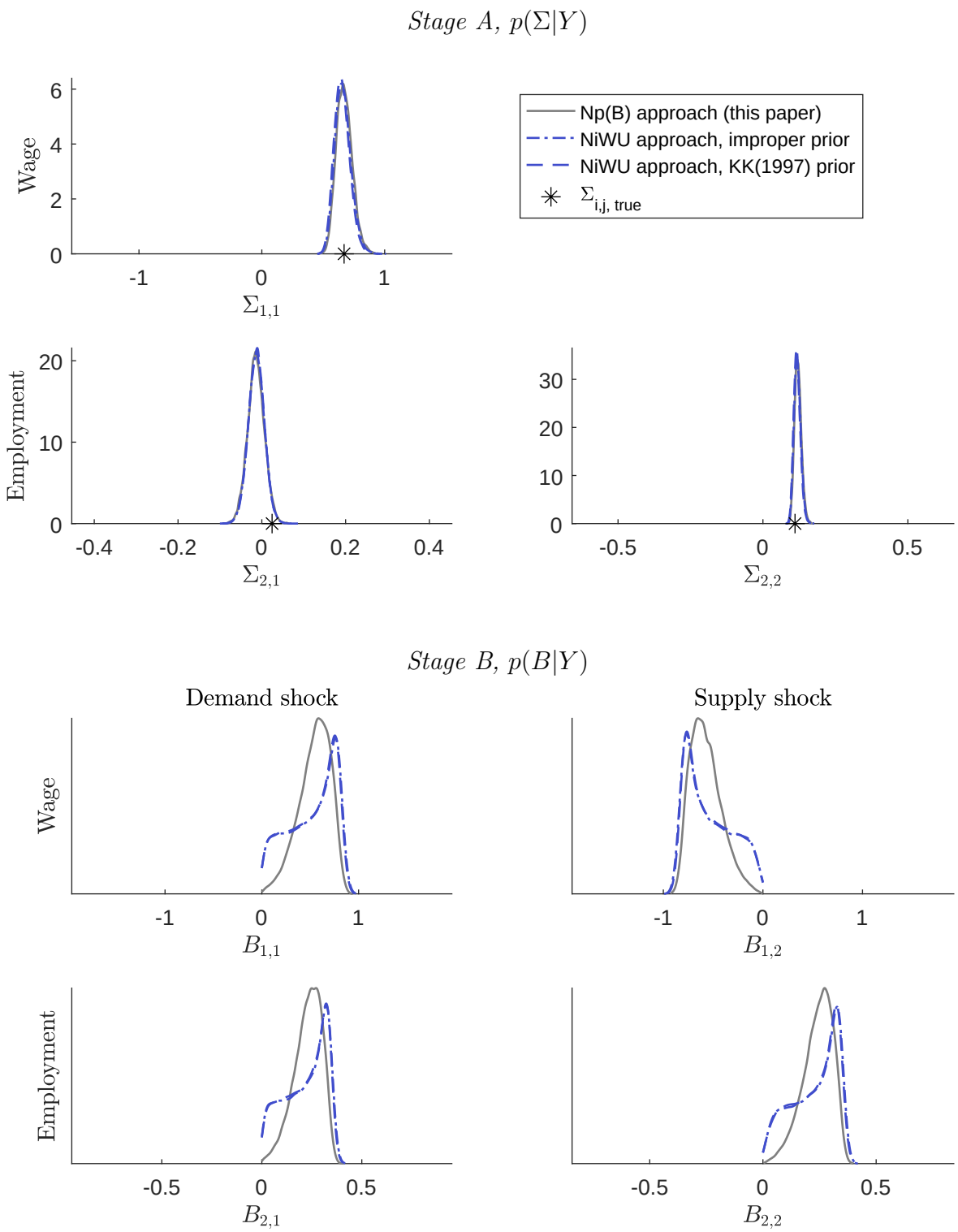


Figure 2.18: Comparing our approach to the NiWU approach, posteriors for $T = 480$

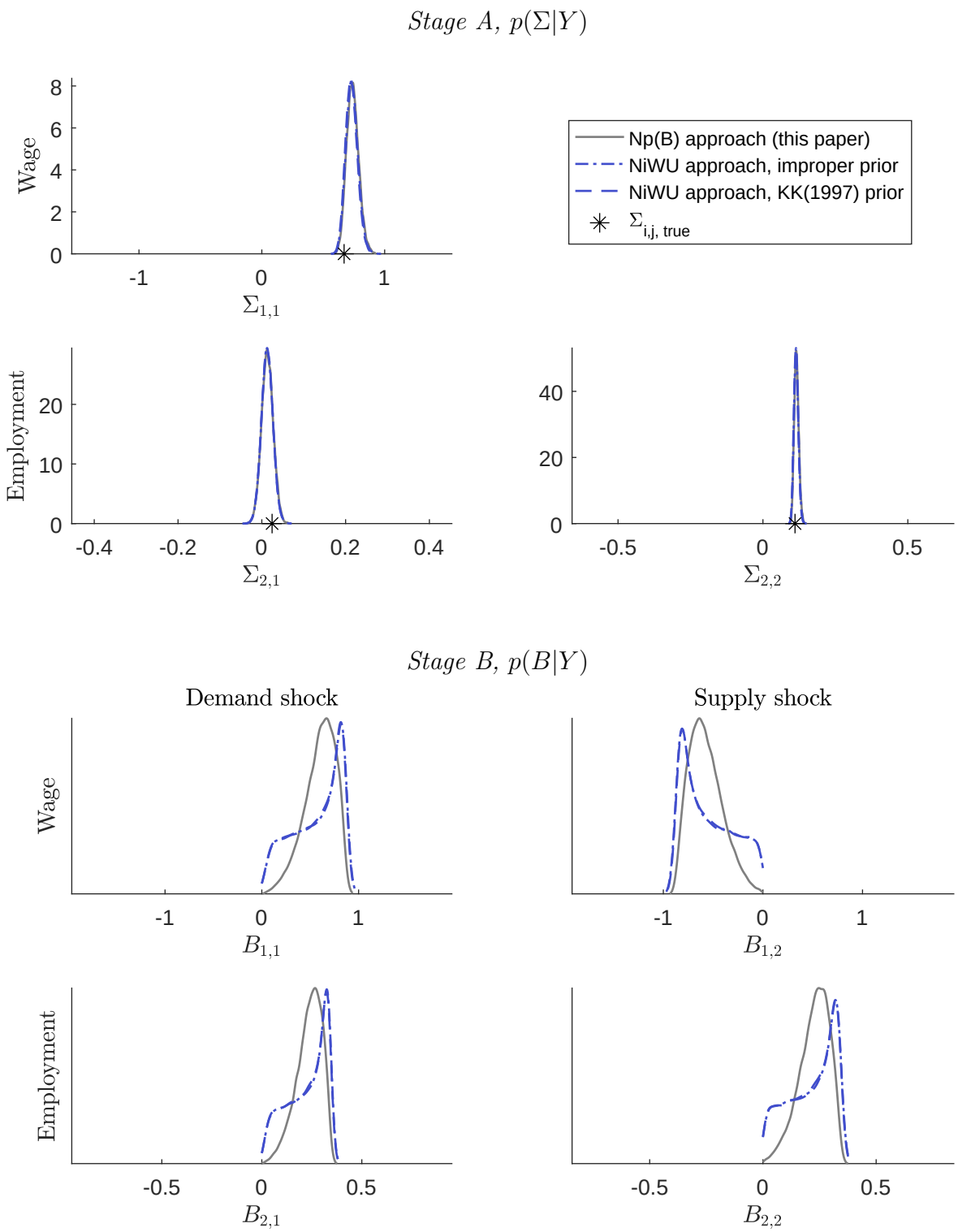


Figure 2.19: Comparing our approach to the NiWU approach, posteriors asymptotically

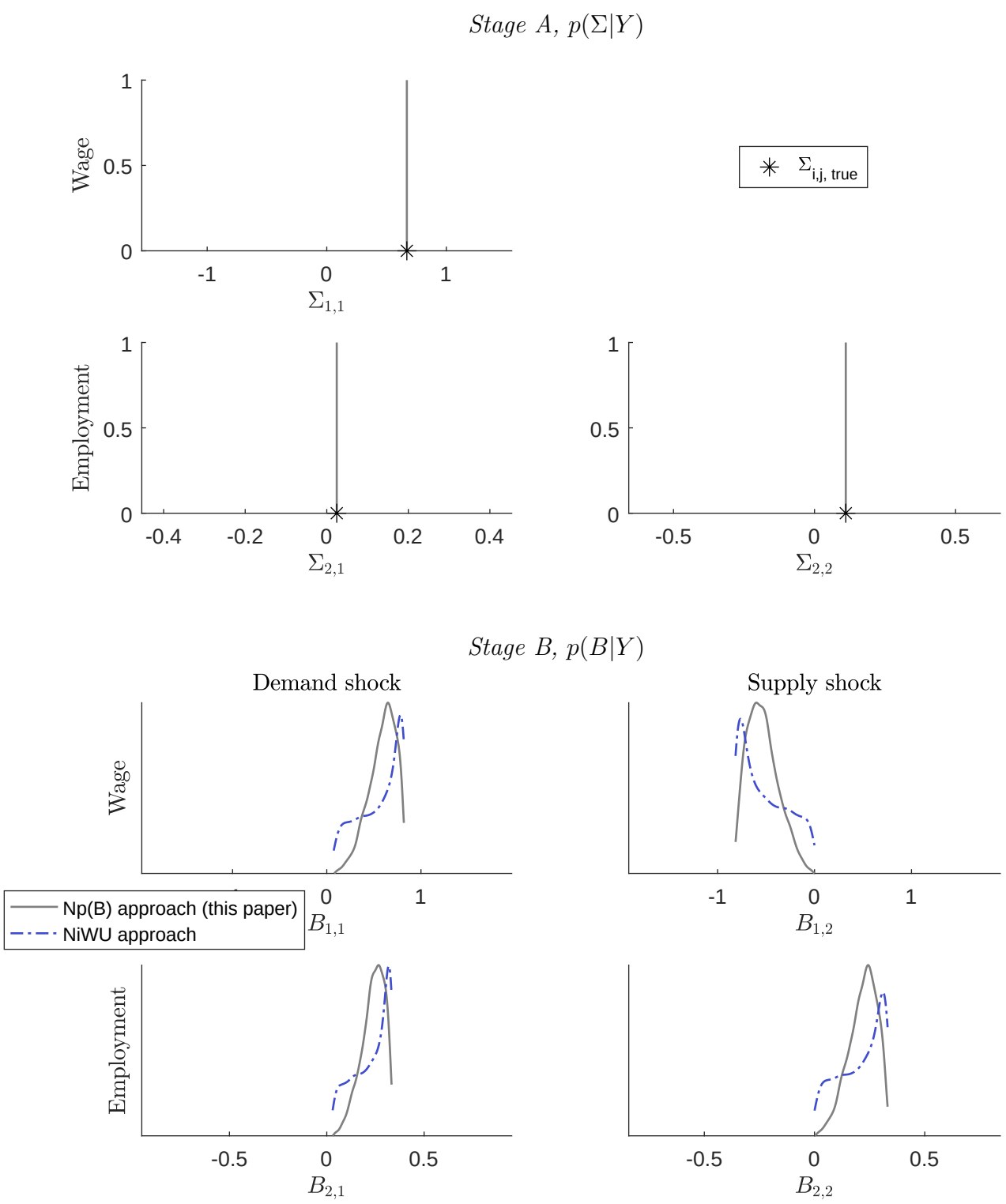
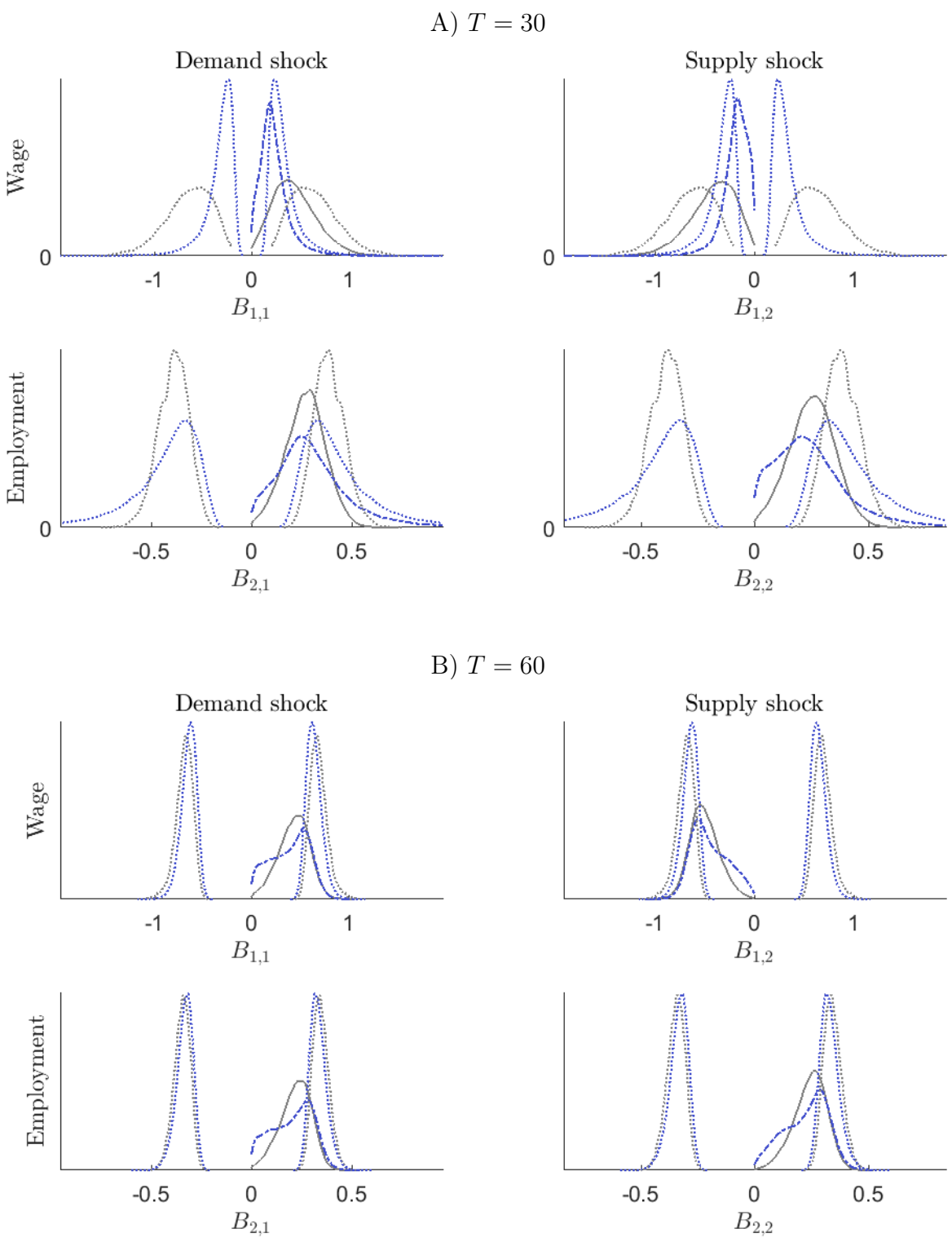
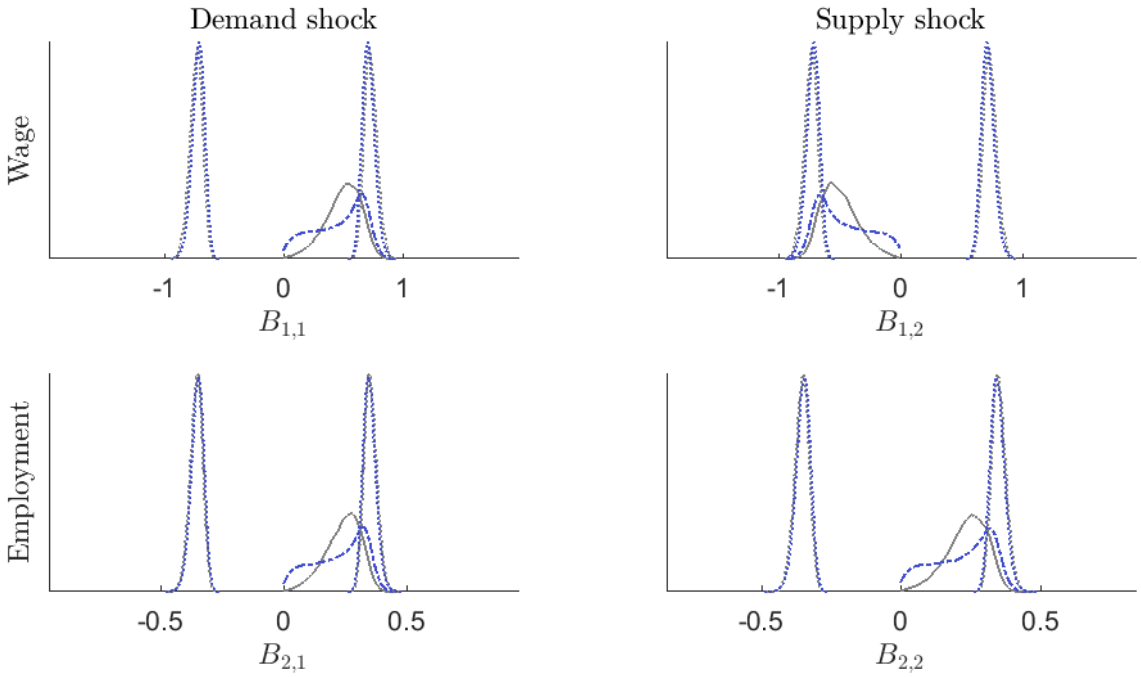


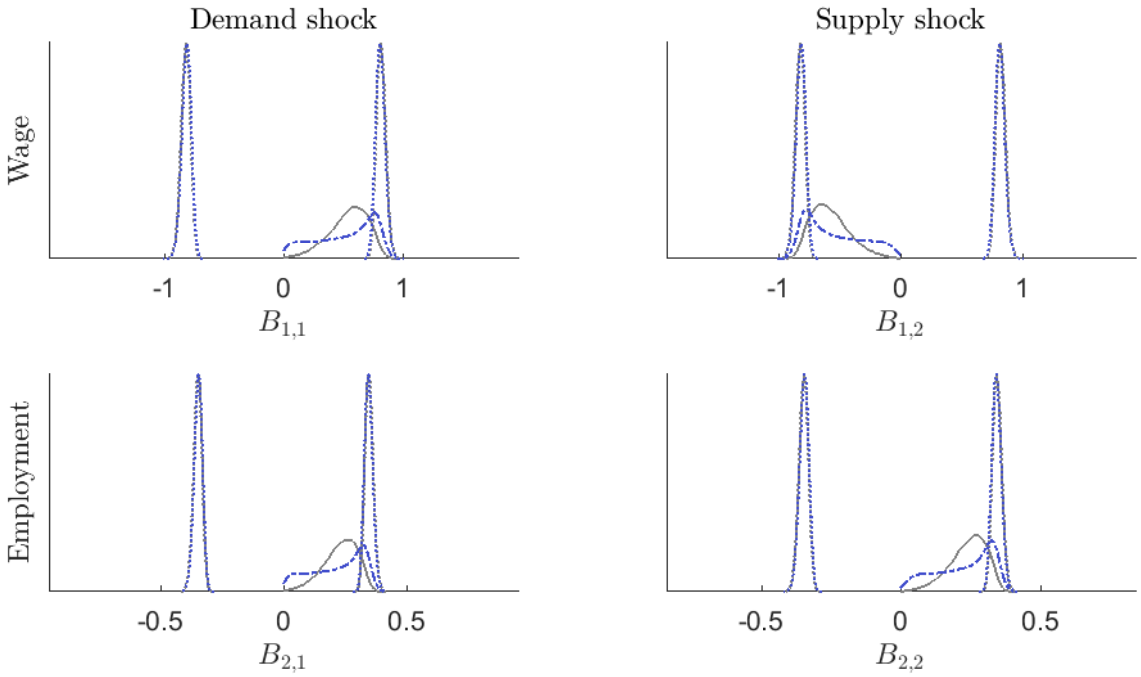
Figure 2.20: Posterior distribution of the estimated bounds $\pm\Sigma_{i,i}^{0.5}$ associated with $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$



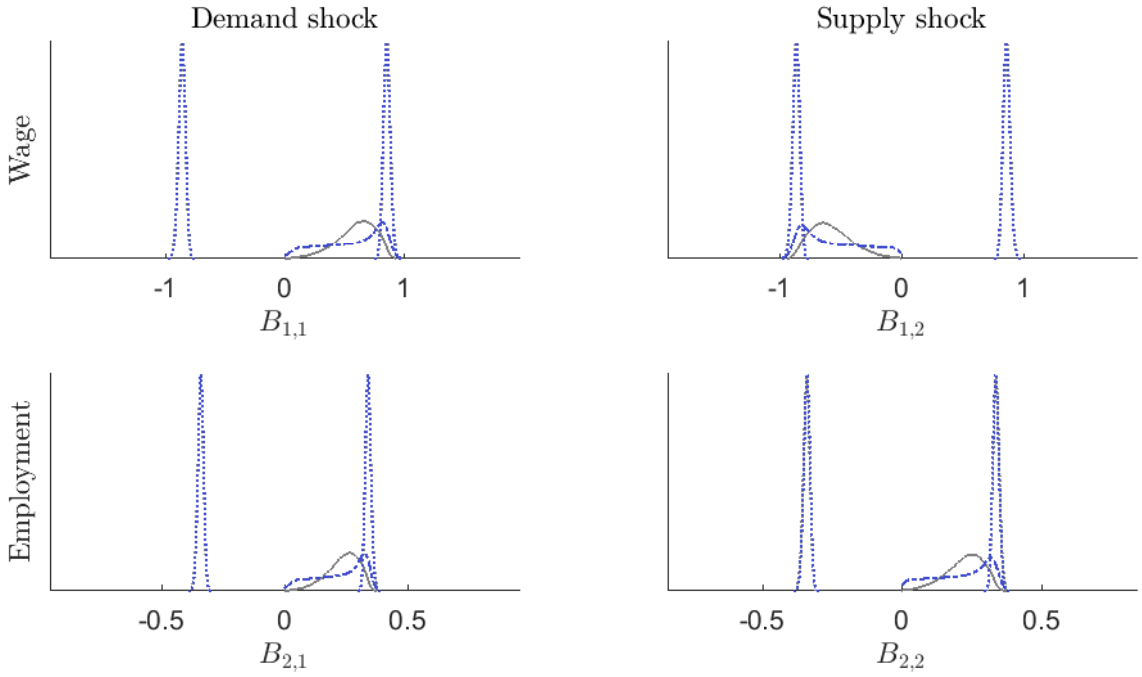
C) $T = 120$


Demand shock
Supply shock
Wage
Employment
-1
0
1
$B_{1,1}$
$B_{1,2}$
-0.5
0
0.5
$B_{2,1}$
$B_{2,2}$

D) $T = 240$


Demand shock
Supply shock
Wage
Employment
-1
0
1
$B_{1,1}$
$B_{1,2}$
-0.5
0
0.5
$B_{2,1}$
$B_{2,2}$

E) $T = 480$



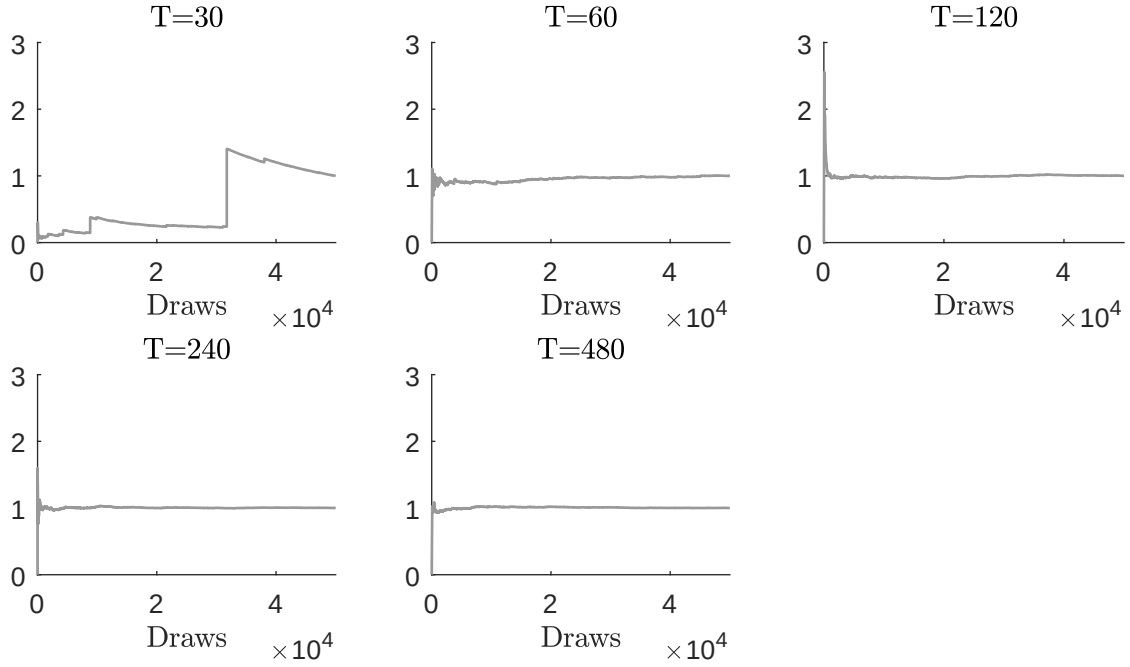
Note: The grey solid line refers to the distribution $p(B|Y)_{Np(B)}$, the dashed blue line refers to the distribution $p(B|Y)_{NiWU}$, the blue and grey dotted lines refer to the distributions of the upper bounds $\pm\Sigma_{i,i}^{0.5}$ corresponding to $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$, respecitvely. The distributions associated with the NiWU approach refer to the improper prior specification.

Table 2.7: Diagnostics on the importance weights, tests

	u	0.5N	0.6N	0.7N	0.9N	0.99N
$T = 30$						
	Wald	-73.00	-60.49	-44.09	22.77	-0.29
	Score	-6.00	-5.94	-5.00	2.75	-0.16
	LR	NA	NA	NA	NA	NA
$T = 60$						
	Wald	-39.96	-34.09	-27.83	-13.27	-2.97
	Score	-13.45	-11.76	-9.98	-5.80	-1.99
	LR	0.00	0.00	0.00	0.00	0.00
$T = 120$						
	Wald	-47.91	-42.59	-36.44	-19.71	-5.36
	Score	-19.08	-16.09	-13.75	-7.51	-2.52
	LR	0.00	0.00	0.00	0.00	0.00
$T = 240$						
	Wald	-53.43	-49.18	-43.69	-26.60	-9.02
	Score	-21.84	-18.86	-16.27	-8.37	-2.78
	LR	0.00	0.00	0.00	0.00	0.00
$T = 480$						
	Wald	-56.71	-53.08	-48.01	-30.60	-10.45
	Score	-23.15	-20.04	-16.76	-9.21	-2.56
	LR	0.00	0.00	0.00	0.00	0.00

Note: Reported are the test statistics. The null hypothesis implies finite weight variance. The corresponding critical values above which the null hypothesis is rejected are 1.65 for the Wald test, 1.65 for the score test and 2.68 for the LR test (5% significance level). The corresponding p-values are close to 0 for $T = 30$ and close to 1 for all other cases.

Figure 2.21: Diagnostics on the importance weights, graphical assessment



Note: The graph shows the recursive variance $\{v_i\}_{i=1}^{N}$, where $v_i = var(w_{1:i})$ computed using de-meaned and standardized weights.

Table 2.8: Convergence diagnostics for the application in section 2.3 of the paper, $T = 30$

parameter	1	2	3	4	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01					
stage 1	0.82	0.33	0.22	0.50	
stage 2	0.38	0.86	0.02	0.09	
stage 3	0.27	0.01	0.92	0.14	
stage 4	0.24	0.05	0.30	0.01	
stage 5	0.72	0.02	0.16	0.12	
stage 6	0.47	0.45	0.41	0.91	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (64,000)					
stage 1	25,457	27,159	30,127	27,500	
stage 2	26,442	26,803	28,344	26,955	
stage 3	24,578	26,511	32,887	25,469	
stage 4	22,904	25,974	32,659	23,816	
stage 5	22,663	21,714	30,463	23,272	
stage 6	24,546	23,686	27,441	23,240	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$					
stage 1	1.003	1.002	1.002	1.002	
stage 2	1.003	1.001	1.002	1.002	
stage 3	1.002	1.002	1.001	1.003	
stage 4	1.002	1.002	1.001	1.003	
stage 5	1.005	1.002	1.002	1.002	
stage 6	1.001	1.001	1.001	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$					
stage 1					1.003
stage 2					1.004
stage 3					1.005
stage 4					1.004
stage 5					1.008
stage 6					1.002

Table 2.9: Convergence diagnostics for the application in section 2.3 of the paper, $T = 60$

parameter	1	2	3	4	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01					
stage 1	0.35	0.93	0.09	0.60	
stage 2	0.07	0.46	0.57	0.17	
stage 3	0.20	0.84	0.74	0.33	
stage 4	0.50	0.50	0.25	0.21	
stage 5	0.26	0.59	0.49	0.60	
stage 6	0.86	0.48	0.96	0.99	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (64,000)					
stage 1	26,633	29,931	20,179	29,286	
stage 2	29,912	27,584	21,278	30,639	
stage 3	27,370	29,397	20,753	30,145	
stage 4	24,342	27,417	22,054	29,624	
stage 5	27,668	24,618	22,783	29,758	
stage 6	25,252	24,410	21,994	24,854	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$					
stage 1	1.001	1.001	1.001	1.000	
stage 2	1.002	1.002	1.002	1.002	
stage 3	1.001	1.000	1.001	1.001	
stage 4	1.000	1.001	1.001	1.001	
stage 5	1.001	1.000	1.001	1.001	
stage 6	1.003	1.002	1.002	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$					
stage 1					1.003
stage 2					1.002
stage 3					1.003
stage 4					1.003
stage 5					1.002
stage 6					1.003

Table 2.10: Convergence diagnostics for the application in section 2.3 of the paper, $T = 120$

parameter	1	2	3	4	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01					
stage 1	0.90	0.84	0.90	0.89	
stage 2	0.64	0.34	0.87	0.66	
stage 3	0.81	0.45	0.57	0.77	
stage 4	0.64	0.18	0.21	0.50	
stage 5	0.25	0.29	0.89	0.67	
stage 6	0.55	0.48	0.42	0.71	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (64,000)					
stage 1	38,177	38,540	24,618	45,268	
stage 2	38,229	37,637	26,218	39,798	
stage 3	38,184	36,551	23,368	37,841	
stage 4	40,124	43,235	22,783	37,114	
stage 5	33,015	37,353	19,773	31,707	
stage 6	30,682	33,353	20,706	31,242	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$					
stage 1	1.001	1.001	1.001	1.001	
stage 2	1.001	1.001	1.001	1.002	
stage 3	1.001	1.001	1.001	1.001	
stage 4	1.001	1.001	1.001	1.001	
stage 5	1.001	1.000	1.000	1.001	
stage 6	1.001	1.001	1.001	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$					
stage 1					1.002
stage 2					1.005
stage 3					1.002
stage 4					1.004
stage 5					1.003
stage 6					1.002

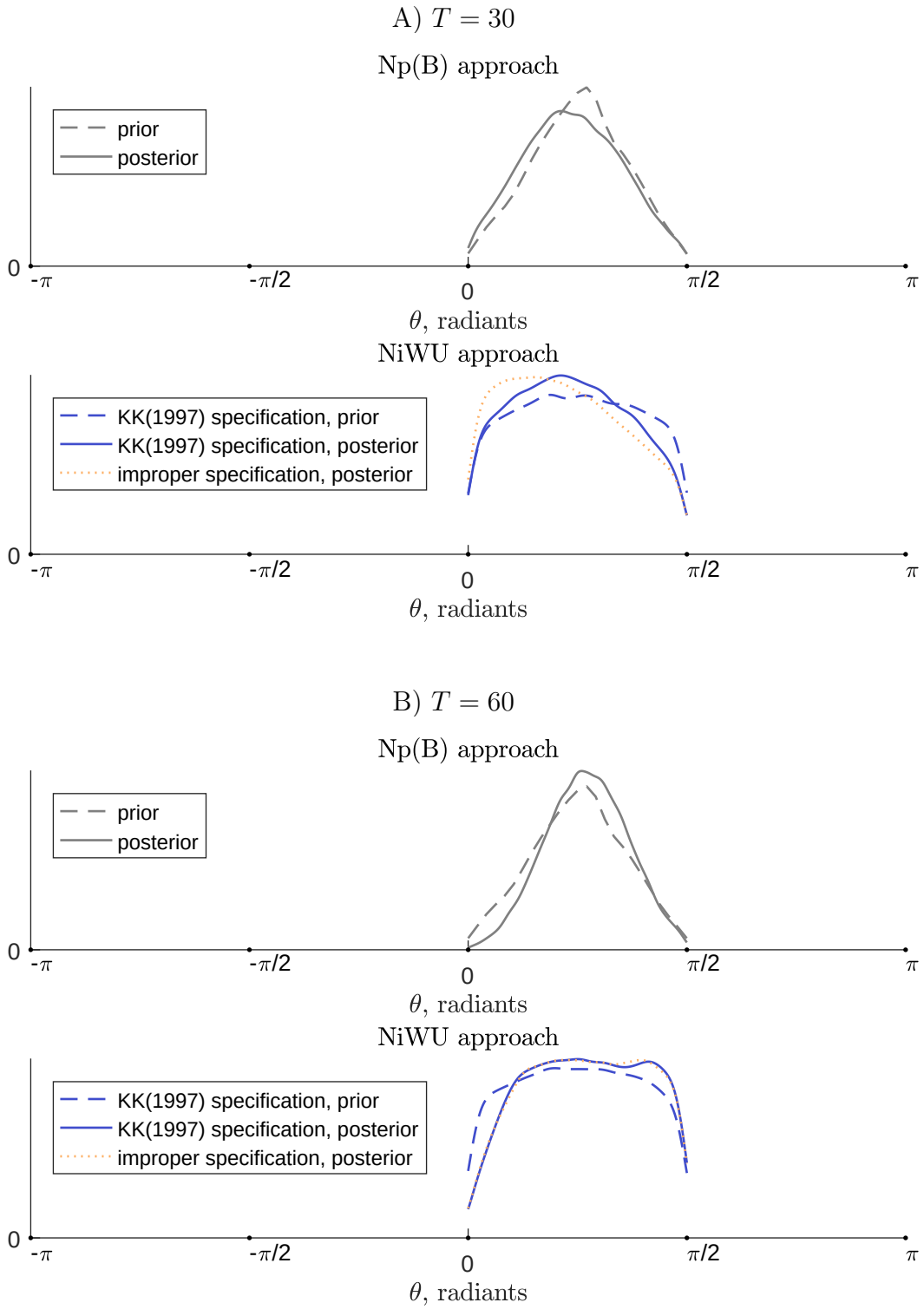
Table 2.11: Convergence diagnostics for the application in section 2.3 of the paper, $T = 240$

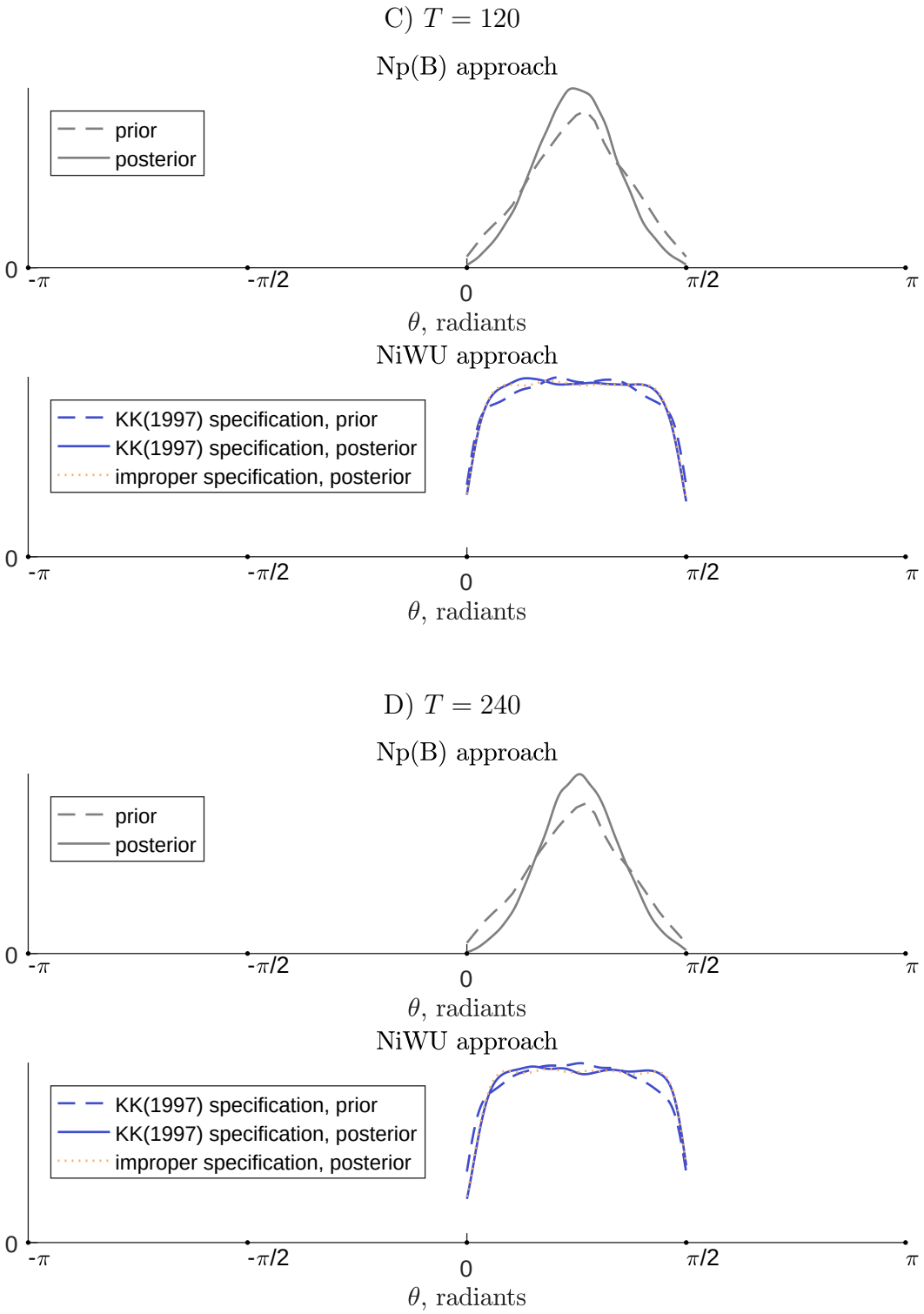
parameter	1	2	3	4	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01					
stage 1	0.28	0.11	0.13	0.35	
stage 2	0.15	0.21	0.27	0.22	
stage 3	0.85	0.91	0.75	0.74	
stage 4	0.74	0.62	0.74	0.68	
stage 5	0.96	0.57	0.96	0.64	
stage 6	0.35	0.18	0.19	0.25	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (64,000)					
stage 1	37,890	44,770	27,980	39,878	
stage 2	43,621	40,519	27,826	41,374	
stage 3	40,011	42,262	27,205	40,226	
stage 4	42,403	40,688	24,506	40,909	
stage 5	39,060	38,904	24,144	37,694	
stage 6	38,379	35,270	20,960	37,987	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$					
stage 1	1.001	1.002	1.002	1.001	
stage 2	1.003	1.003	1.004	1.002	
stage 3	1.003	1.003	1.004	1.003	
stage 4	1.001	1.003	1.003	1.001	
stage 5	1.001	1.002	1.002	1.001	
stage 6	1.001	1.001	1.001	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$					
stage 1					1.004
stage 2					1.005
stage 3					1.006
stage 4					1.004
stage 5					1.003
stage 6					1.001

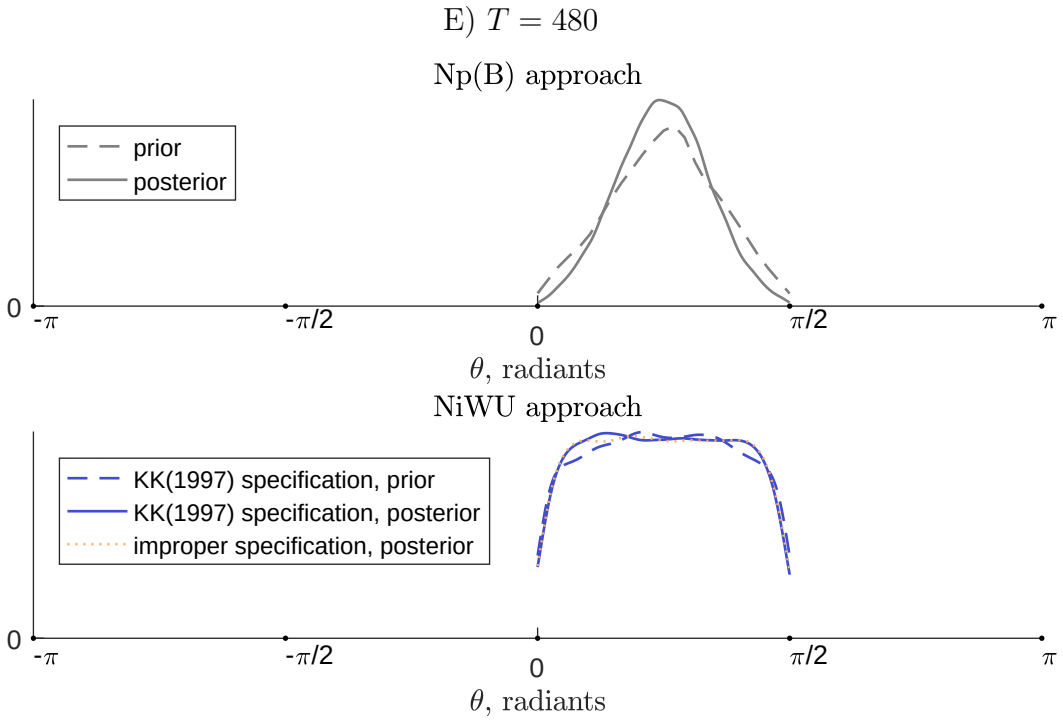
Table 2.12: Convergence diagnostics for the application in section 2.3 of the paper, $T = 480$

parameter	1	2	3	4	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01					
stage 1	0.15	0.24	0.18	0.14	
stage 2	0.69	0.82	0.76	0.79	
stage 3	0.81	0.92	0.84	0.97	
stage 4	0.19	0.13	0.17	0.38	
stage 5	0.06	0.01	0.01	0.02	
stage 6	0.16	0.09	0.15	0.35	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (64,000)					
stage 1	45,528	49,194	35,838	46,921	
stage 2	47,989	41,589	31,108	44,513	
stage 3	47,795	46,939	32,032	49,132	
stage 4	43,952	44,744	29,345	41,305	
stage 5	45,518	45,321	28,359	44,232	
stage 6	41,568	45,321	27,007	39,536	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$					
stage 1	1.003	1.002	1.003	1.003	
stage 2	1.003	1.003	1.003	1.004	
stage 3	1.002	1.002	1.002	1.001	
stage 4	1.001	1.001	1.001	1.001	
stage 5	1.003	1.004	1.004	1.003	
stage 6	1.001	1.001	1.001	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$					
stage 1					1.003
stage 2					1.005
stage 3					1.006
stage 4					1.003
stage 5					1.005
stage 6					1.003

Figure 2.22: Marginal distribution on the rotation angle θ implicit in the marginal prior and posterior distributions $p(Q)$ and $p(Q|Y)$




C) $T = 120$
Np(B) approach
prior
posterior
0
-π
-π/2
0
π/2
π
θ, radiants
NiWU approach
KK(1997) specification, prior
KK(1997) specification, posterior
improper specification, posterior
0
-π
-π/2
0
π/2
π
θ, radiants
D) $T = 240$
Np(B) approach
prior
posterior
0
-π
-π/2
0
π/2
π
θ, radiants
NiWU approach
KK(1997) specification, prior
KK(1997) specification, posterior
improper specification, posterior
0
-π
-π/2
0
π/2
π
θ, radiants



Note: The NiWU approach uses the distribution $p(Q|\Sigma)_{NiWU}$, which is the uniform distribution in the matrix parameter space $\mathcal{Q}_\Sigma$. The Np(B) approach uses the distribution $p(Q|\Sigma)_{Np(B)}$ that is implied by the prior $p(B)_{Np(B)}$. The marginal distributions $p(Q)_{NiWU}$, $p(Q|Y)_{NiWU}$, $p(Q)_{Np(B)}$ and $p(Q|Y)_{Np(B)}$ are obtained by marginalizing out Σ using $p(\Sigma)_{NiWU}$, $p(\Sigma|Y)_{NiWU}$, $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{Np(B)}$, respectively. The figure shows the distribution on the rotation angle θ associated with the distributions $p(Q)_{NiWU}$, $p(Q|Y)_{NiWU}$, $p(Q)_{Np(B)}$ and $p(Q|Y)_{Np(B)}$. See also the discussion in footnote 12 of the paper and section 2.A.3.2 of the Appendix.

2.A.6 Additional tables/figures for section 2.4

Figure 2.23: Data, as it enters the model

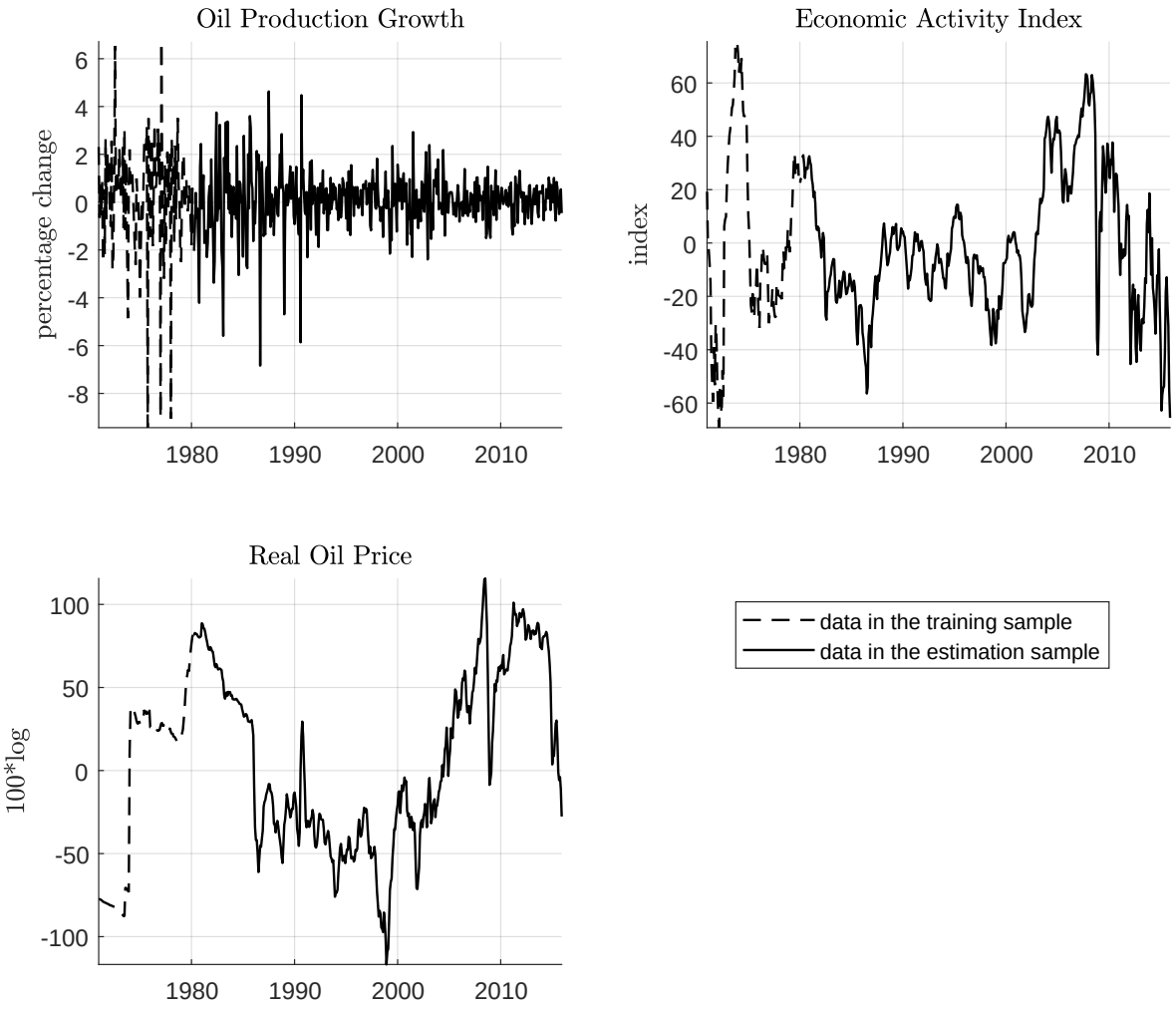
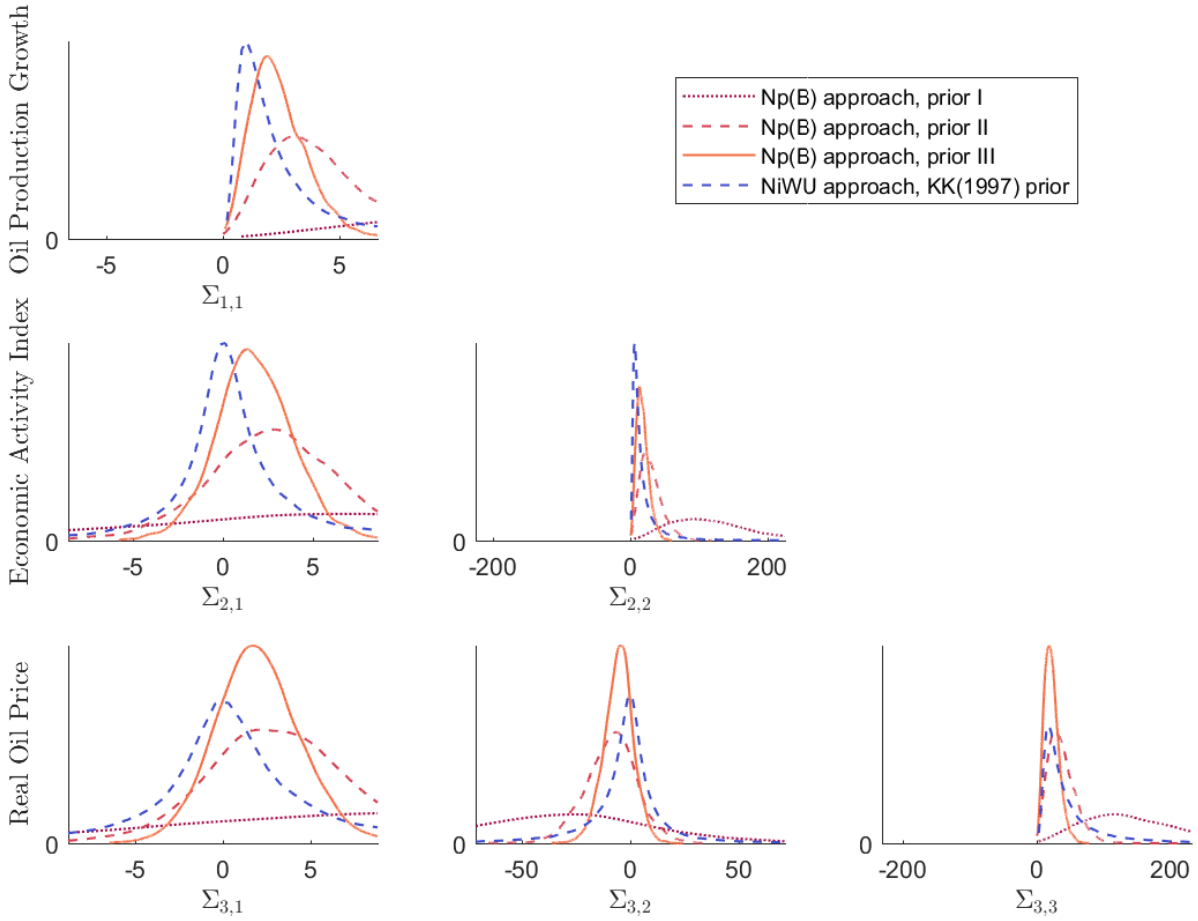
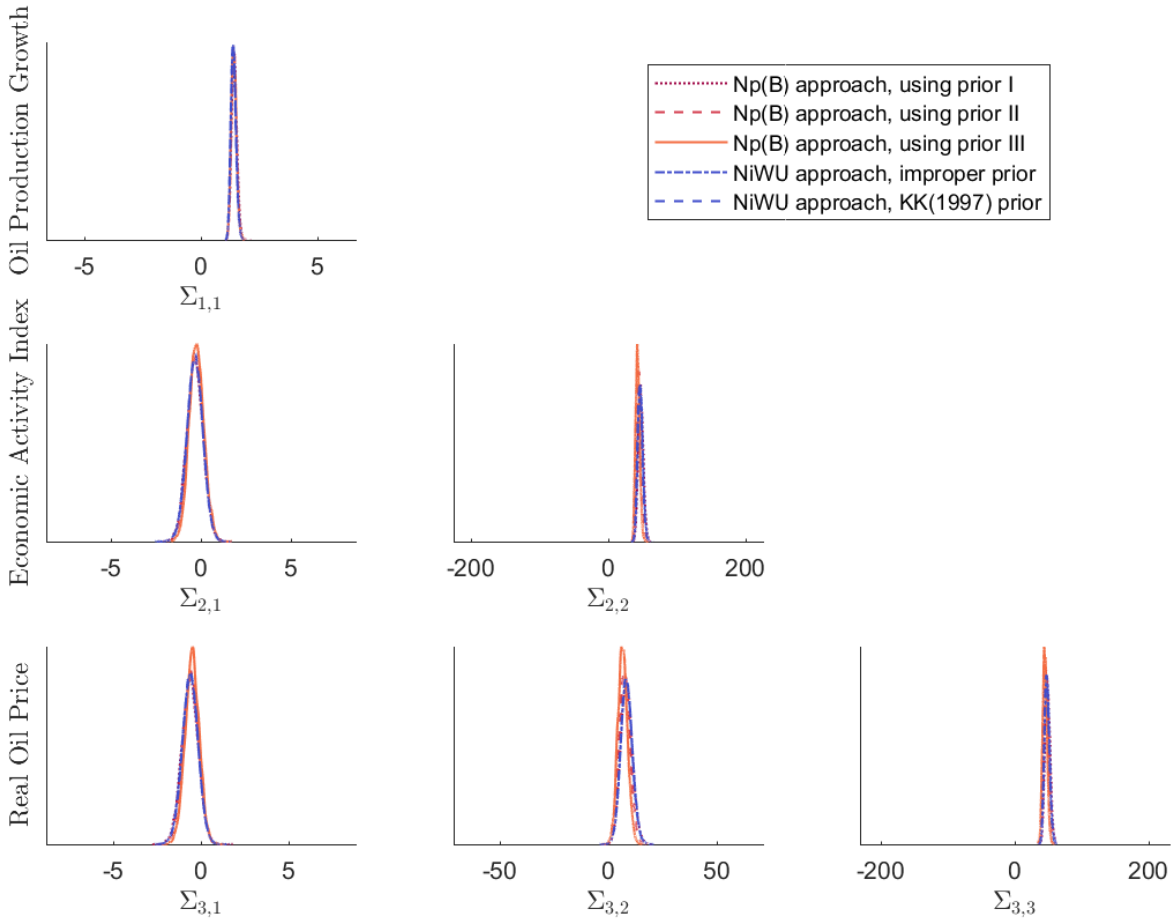


Figure 2.24: $p(\Sigma)_{NiWU}$ and $p(\Sigma)_{Np(B)}$



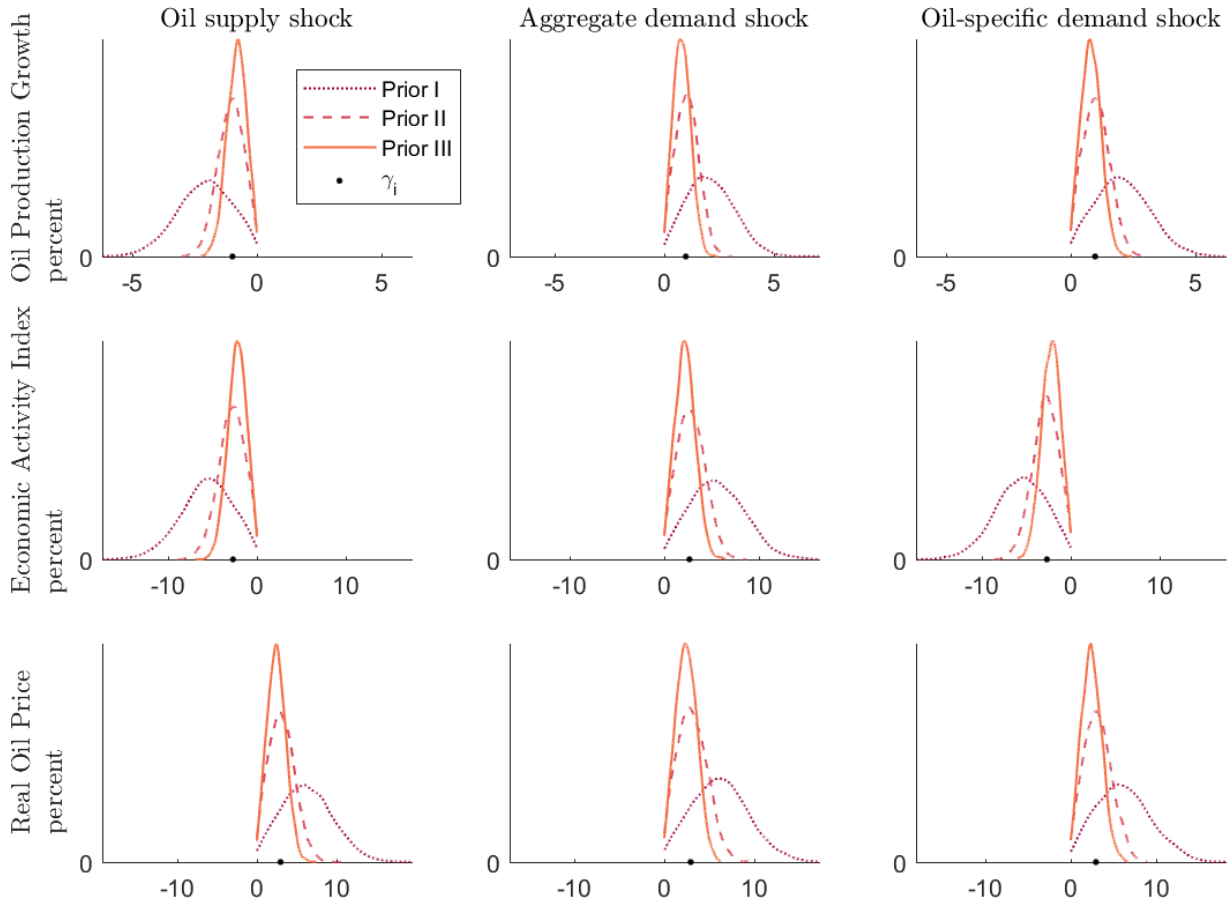
Note: The figure shows the prior distribution $p(\Sigma)$ associated with our approach and with the prior specification by Kadiyala and Karlsson (1997) of the NiWU approach.

Figure 2.25: $p(\Sigma|Y)_{NiWU}$ and $p(\Sigma|Y)_{Np(B)}$



Note: The figure shows the posterior distribution $p(\Sigma|Y)$ associated with our approach and with the NiWU approach.

Figure 2.26: $p(B)_{Np(B)}$ comparing prior specifications for our approach



Note: The figure shows the three alternative prior specifications we use for the prior distribution in our Np(B) approach. See also table 2.3 in the paper.

Table 2.13: Allocation of prior probability mass in our Np(B) approach

Estimated bounds γ_i used to specify the prior $p(B)_{Np(B)}$

$i = 1$: Oil production growth	0.978
$i = 2$: Economic activity index	2.668
$i = 3$: Real oil price	2.953

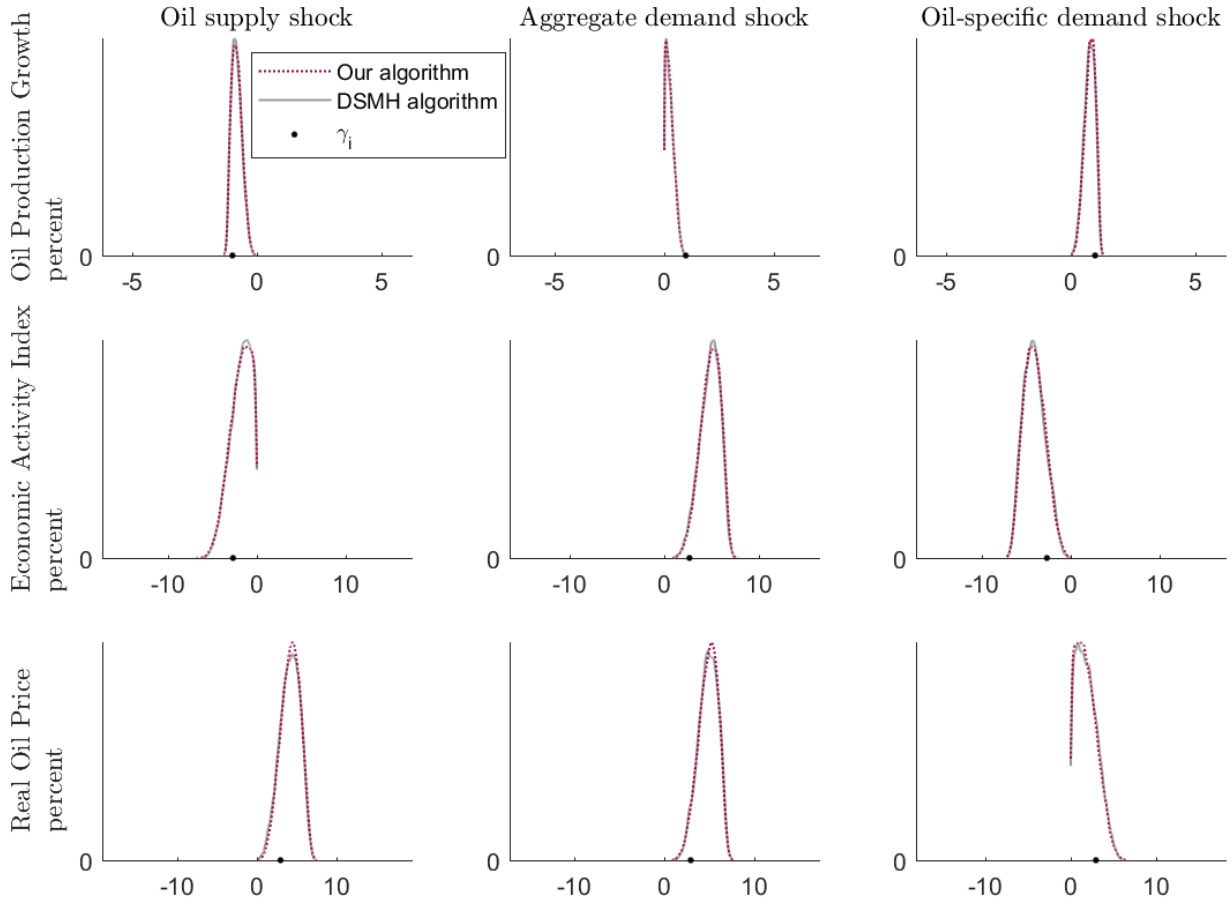
Prior probability mass of the absolute value of $B_{i,j}$ beyond γ_i

prior I			prior II			prior III		
0.84	0.84	0.84	0.52	0.52	0.52	0.33	0.33	0.33
0.84	0.84	0.84	0.53	0.53	0.53	0.33	0.33	0.33
0.84	0.84	0.84	0.53	0.53	0.53	0.33	0.33	0.33

97.5^{th} percentile of the marginal distributions $p(B_{i,j})_{Np(B)}$

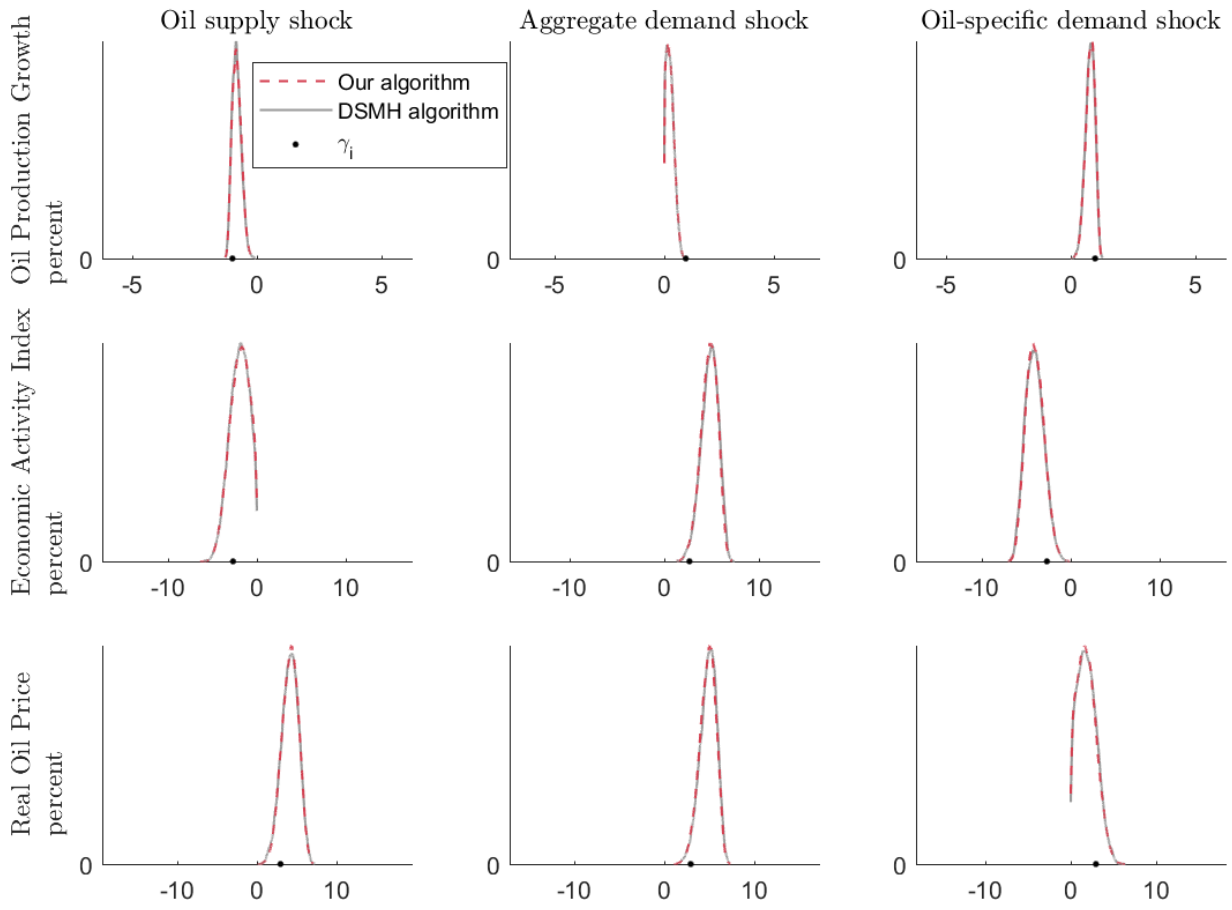
prior I			prior II			prior III		
-4.353	4.341	4.352	-2.124	2.097	2.103	-1.625	1.604	1.620
-11.721	11.492	-11.537	-5.960	5.787	-5.929	-4.349	4.396	-4.372
12.890	12.889	12.868	6.481	6.501	6.487	4.801	4.756	4.887

Figure 2.27: $p(B|Y)_{Np(B)}$, comparing algorithms (prior I)



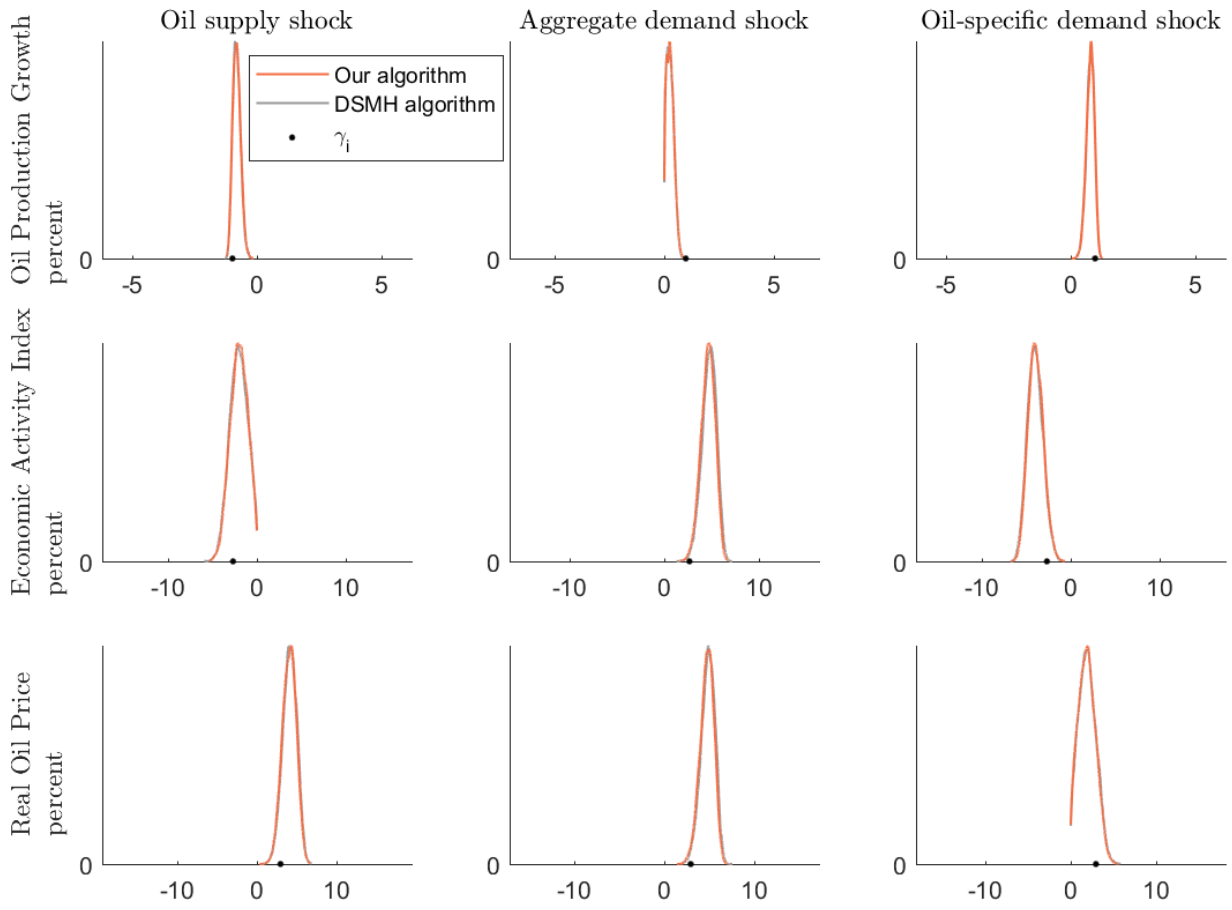
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification A (table 2.3 in the paper), sampled using either our algorithm or the Dynamic Strated Metropolis-Hastings algorithm by Waggoner et al. (2016).

Figure 2.28: $p(B|Y)_{Np(B)}$, comparing algorithms (prior II)



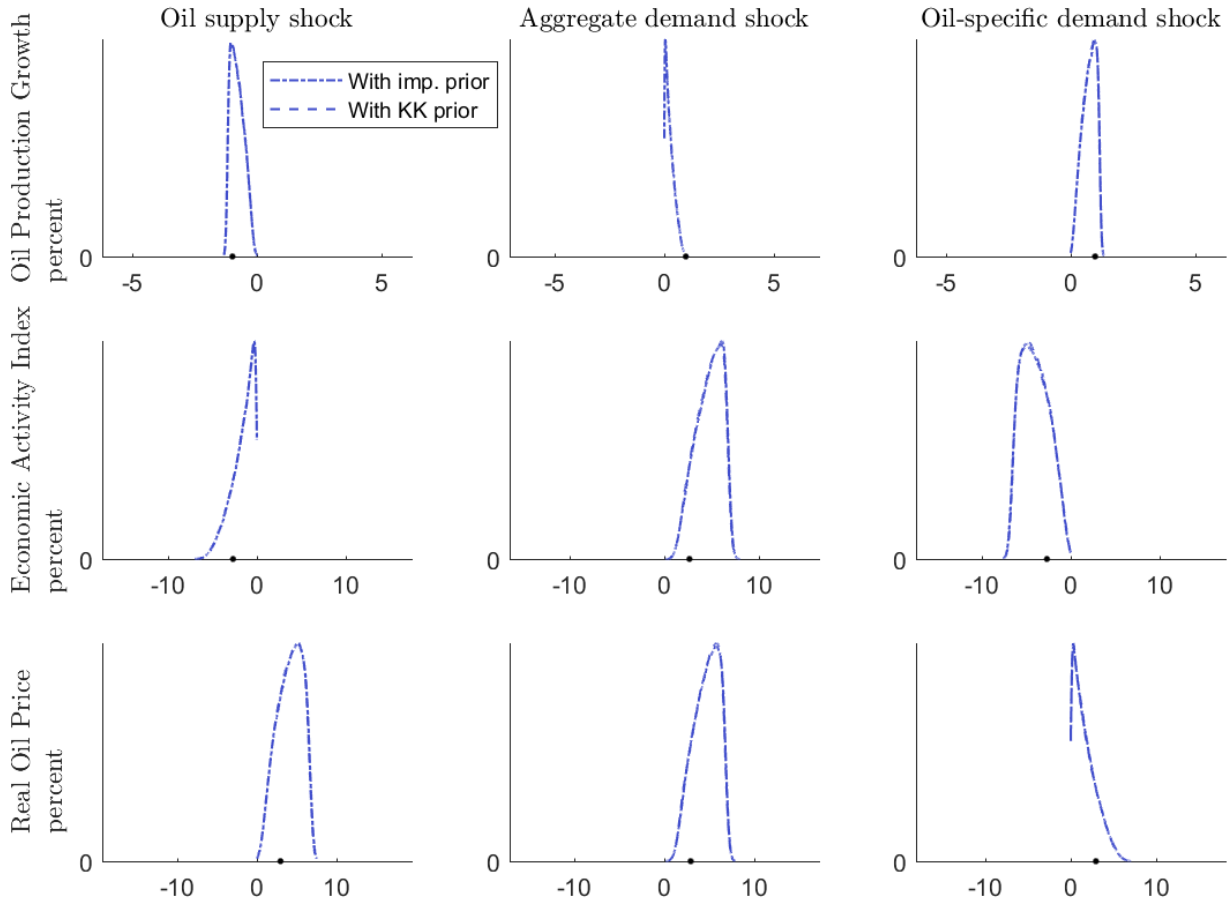
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification B (table 2.3 in the paper), sampled using either our algorithm or the Dynamic Strated Metropolis-Hastings algorithm by Waggoner et al. (2016).

Figure 2.29: $p(B|Y)_{Np(B)}$, comparing algorithms (prior III)



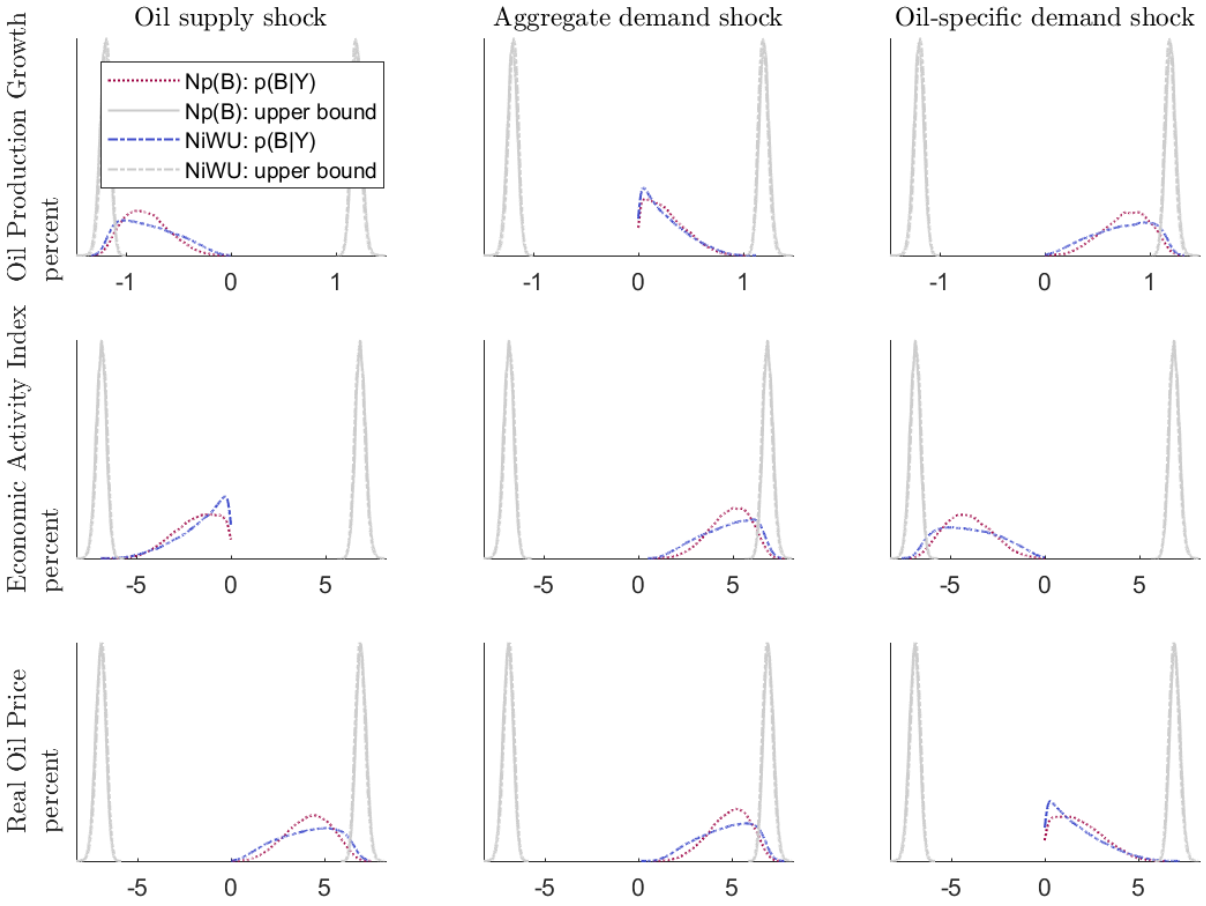
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification C (table 2.3 in the paper), sampled using either our algorithm or the Dynamic Strated Metropolis-Hastings algorithm by Waggoner et al. (2016).

Figure 2.30: $p(B|Y)_{NiWU}$, comparing NiWU approach with improper prior and with prior from Kadiyala and Karlsson (1997)



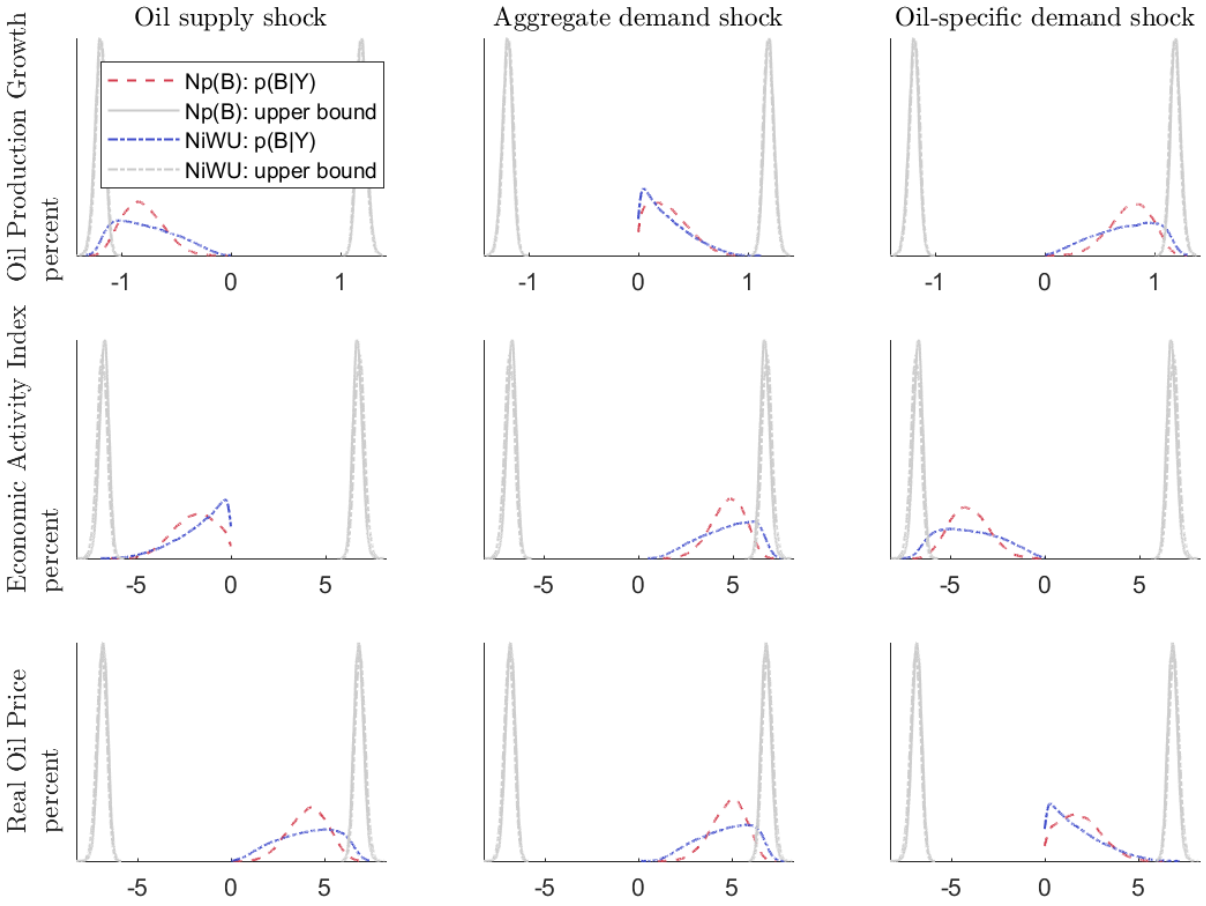
Note: The figure shows the posterior distribution $p(B|Y)_{NiWU}$ when parametrizing the underlying inverse Wishart either with the improper prior or with the prior specification by Kadiyala and Karlsson (1997).

Figure 2.31: $p(B|Y)$, comparing our approach to the NiWU approach (prior I)



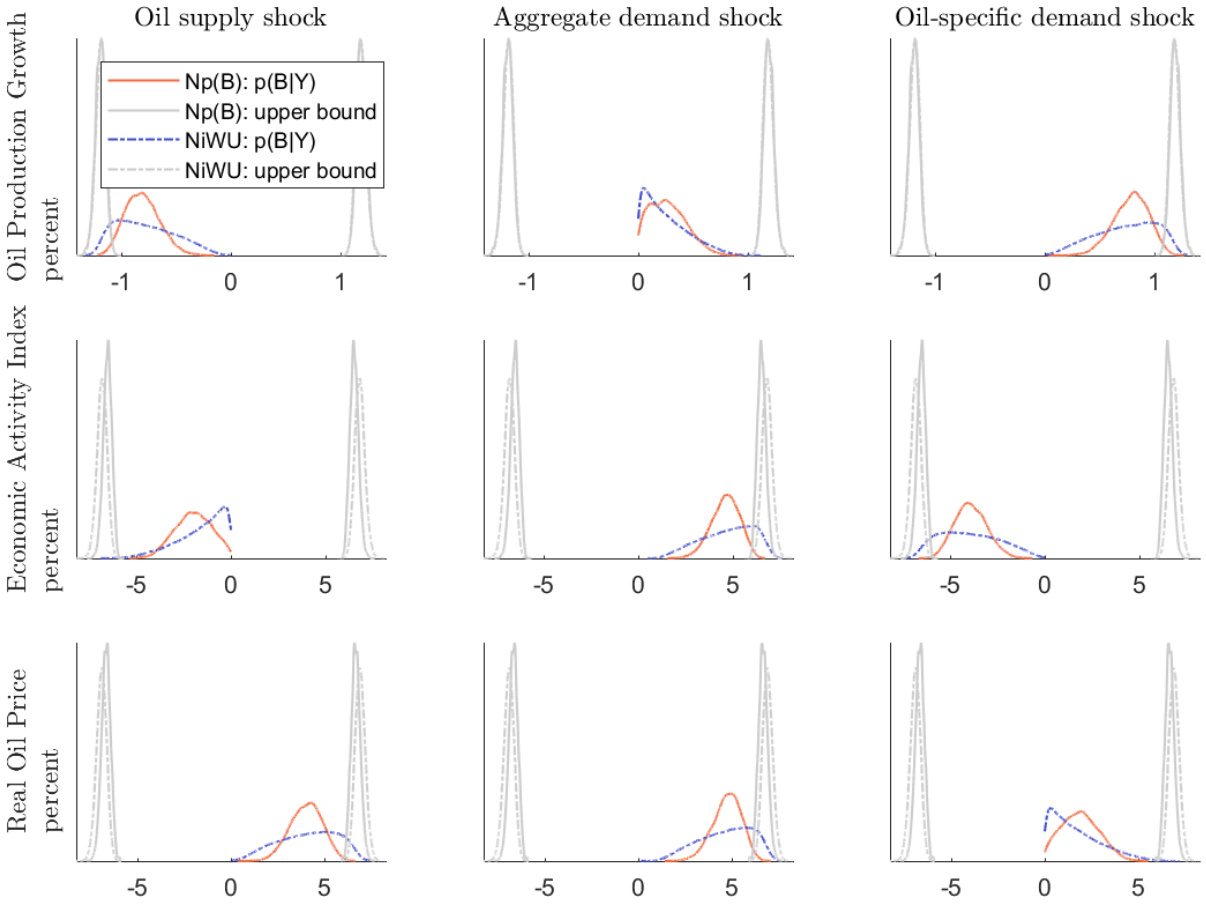
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification I (table 2.3 in the paper) and the posterior distribution $p(B|Y)_{NiWU}$ associated with the improper prior specification. It then shows the distributions of the upper bounds $\pm\Sigma_{i,i}^{0.5}$ (equation (2.11) in the paper) associated with the posterior distributions $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$.

Figure 2.32: $p(B|Y)$, comparing our approach to the NiWU approach (prior II)



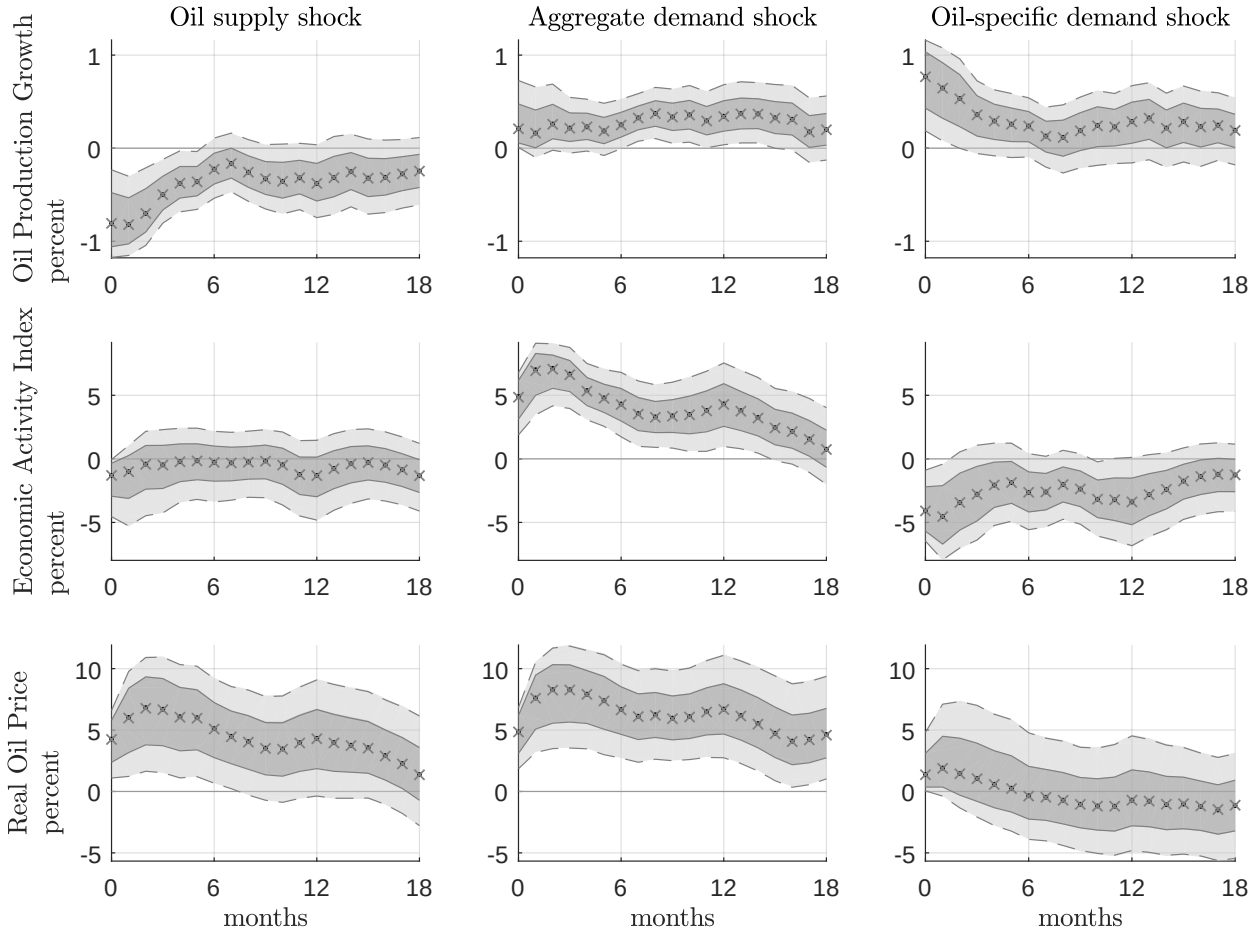
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification II (table 2.3 in the paper) and the posterior distribution $p(B|Y)_{NiWU}$ associated with the improper prior specification. It then shows the distributions of the upper bounds $\pm\Sigma_{i,i}^{0.5}$ (equation (2.11) in the paper) associated with the posterior distributions $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$.

Figure 2.33: $p(B|Y)$, comparing our approach to the NiWU approach (prior III)



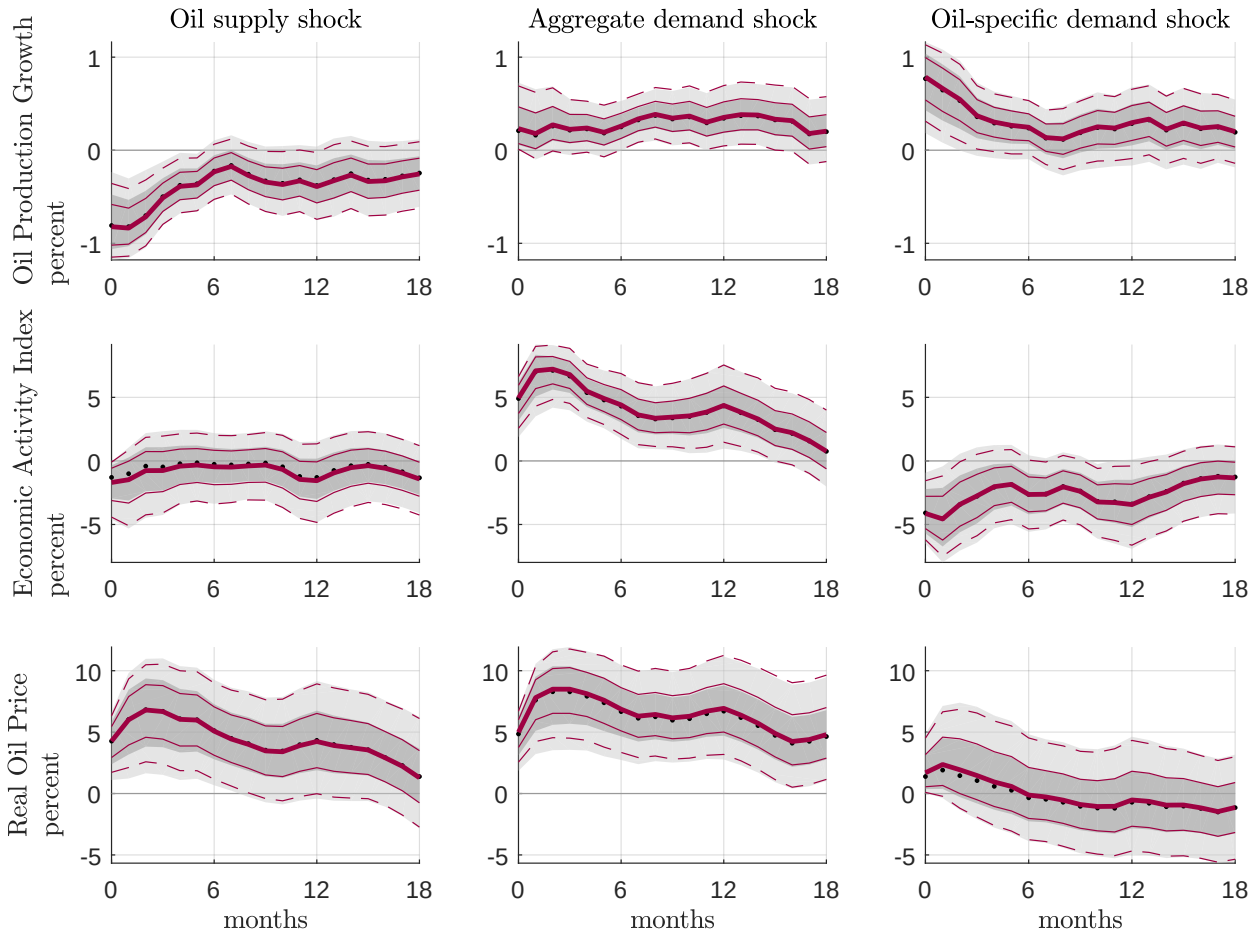
Note: The figure shows the posterior distribution $p(B|Y)_{Np(B)}$ associated with prior specification III (table 2.3 in the paper) and the posterior distribution $p(B|Y)_{NiWU}$ associated with the improper prior specification. It then shows the distributions of the upper bounds $\pm\Sigma_{i,i}^{0.5}$ (equation (2.11) in the paper) associated with the posterior distributions $p(\Sigma|Y)_{Np(B)}$ and $p(\Sigma|Y)_{NiWU}$.

Figure 2.34: IRFs, comparing NiWU approach with improper prior and with prior from Kadiyala and Karlsson (1997)



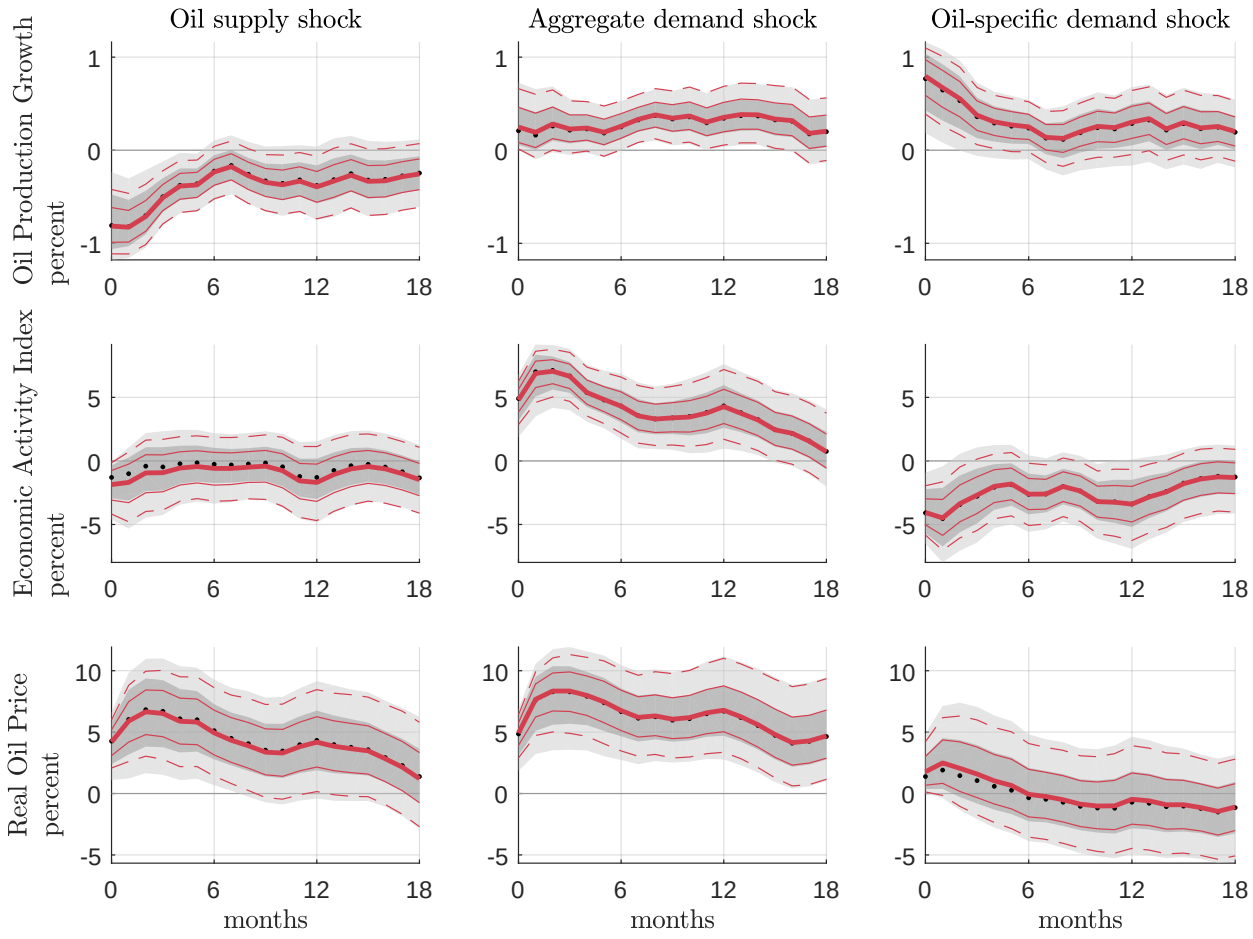
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the inverse Wishart distribution. The x-ed, dashed and solid lines show the same statistics associated with the prior specification of the inverse Wishart by Kadiyala and Karlsson (1997). As in Antolín-Díaz and Rubio-Ramírez (2018), the response of the oil production has been accumulated to the level.

Figure 2.35: IRFs, comparing NiWU and Np(B) (prior I)



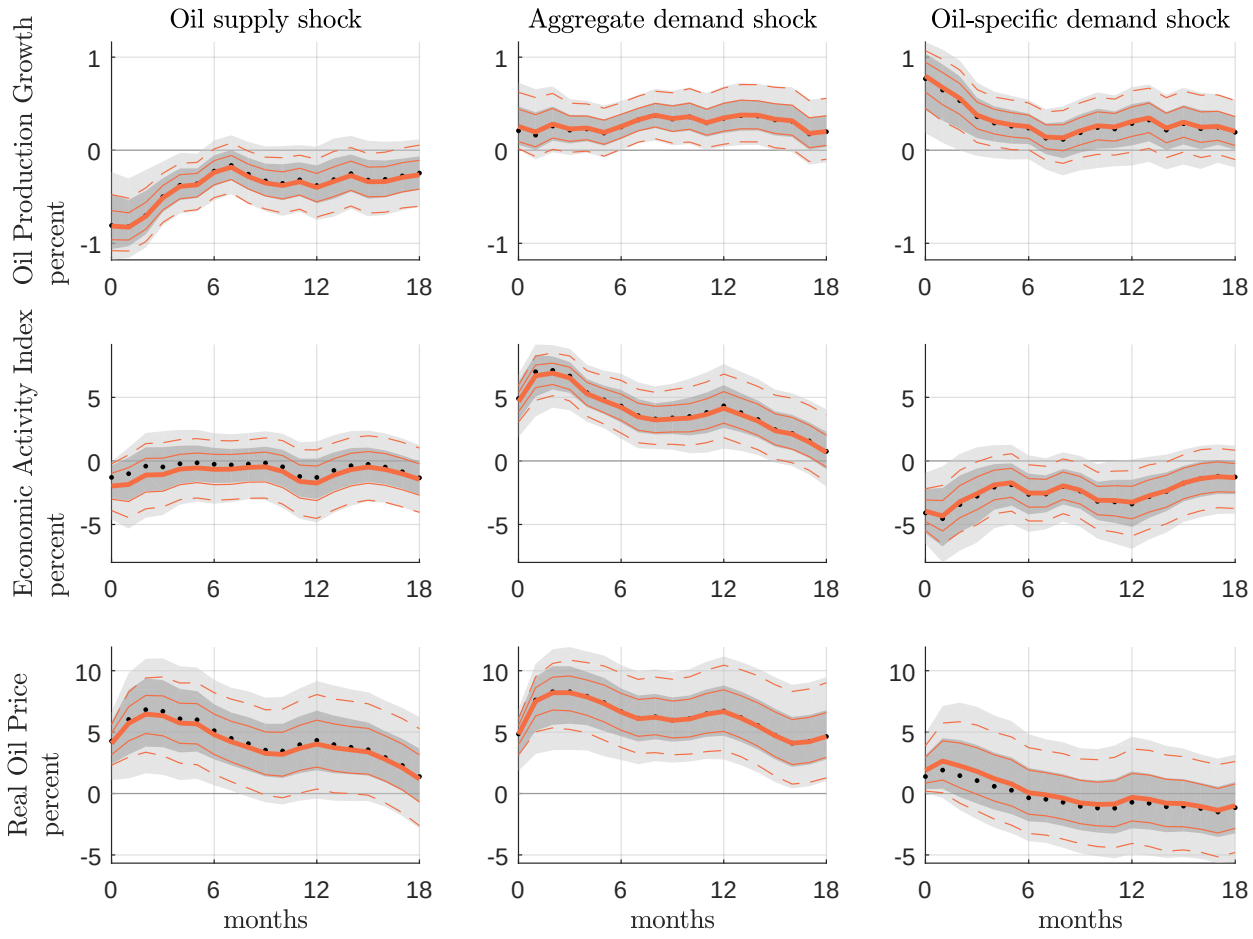
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification I. As in Antolín-Díaz and Rubio-Ramírez (2018), the response of the oil production has been accumulated to the level.

Figure 2.36: IRFs, comparing NiWU and Np(B) (prior II)



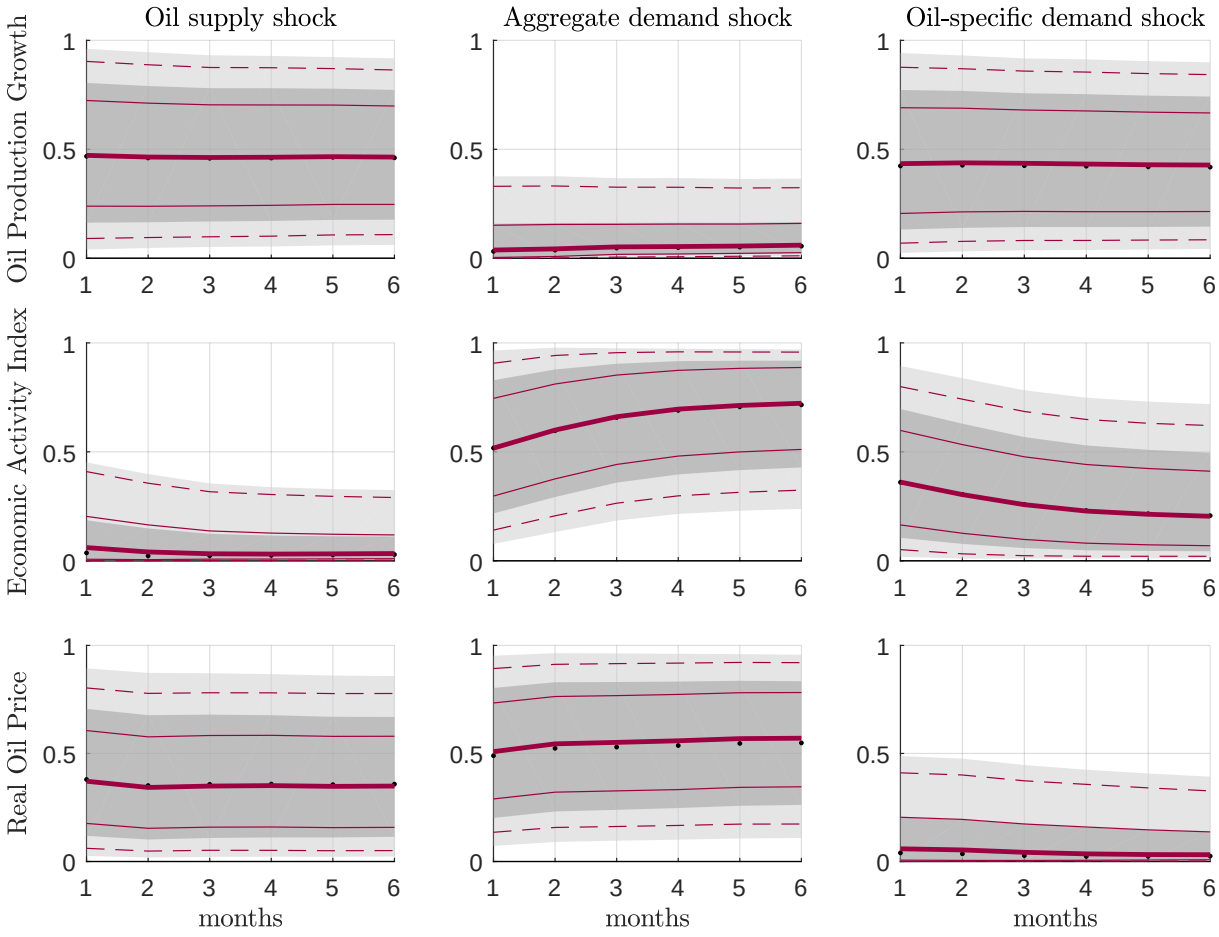
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification II. As in Antolín-Díaz and Rubio-Ramírez (2018), the response of the oil production has been accumulated to the level.

Figure 2.37: IRFs, comparing NiWU and Np(B) (prior III)



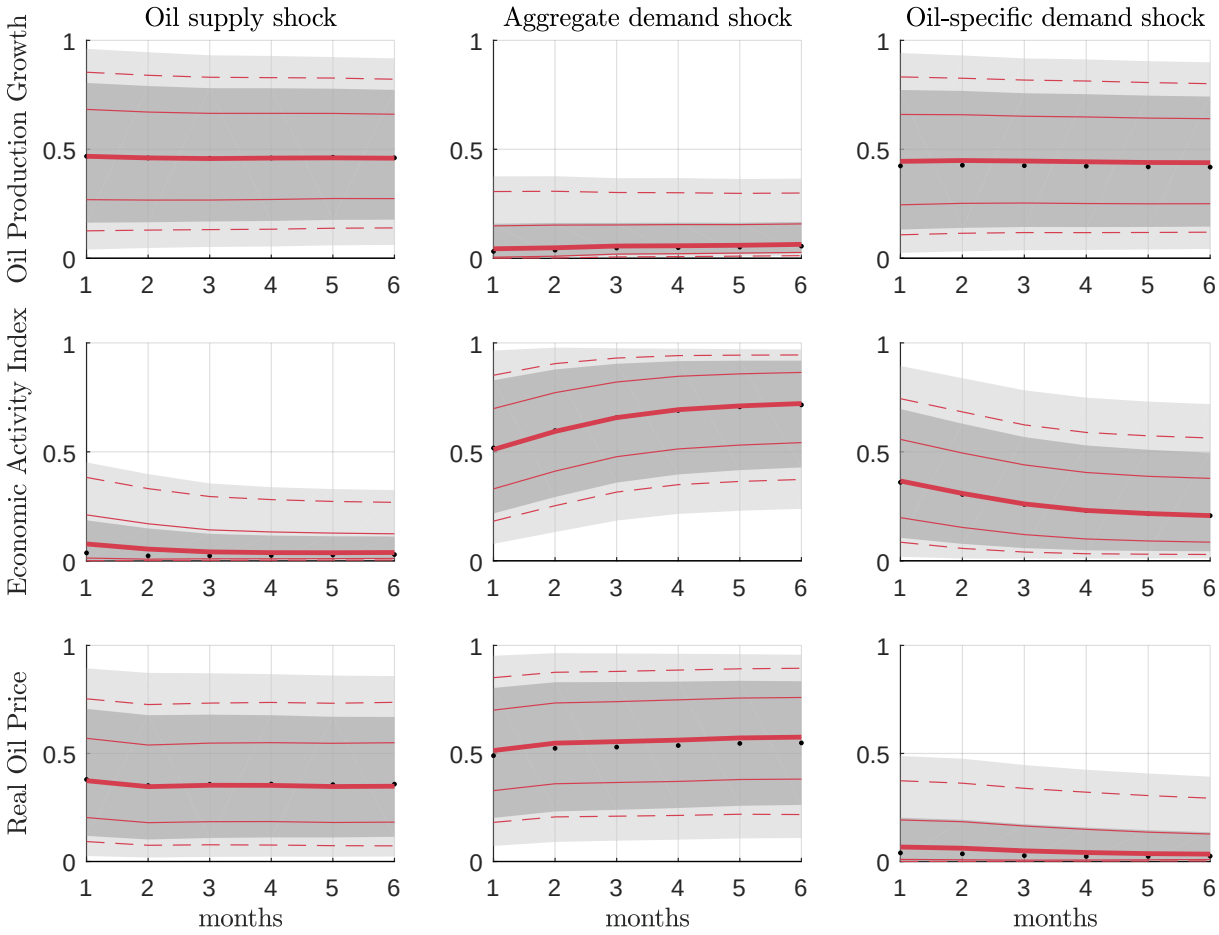
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification III. As in Antolín-Díaz and Rubio-Ramírez (2018), the response of the oil production has been accumulated to the level.

Figure 2.38: FEVD, comparing NiWU and Np(B) (prior I)



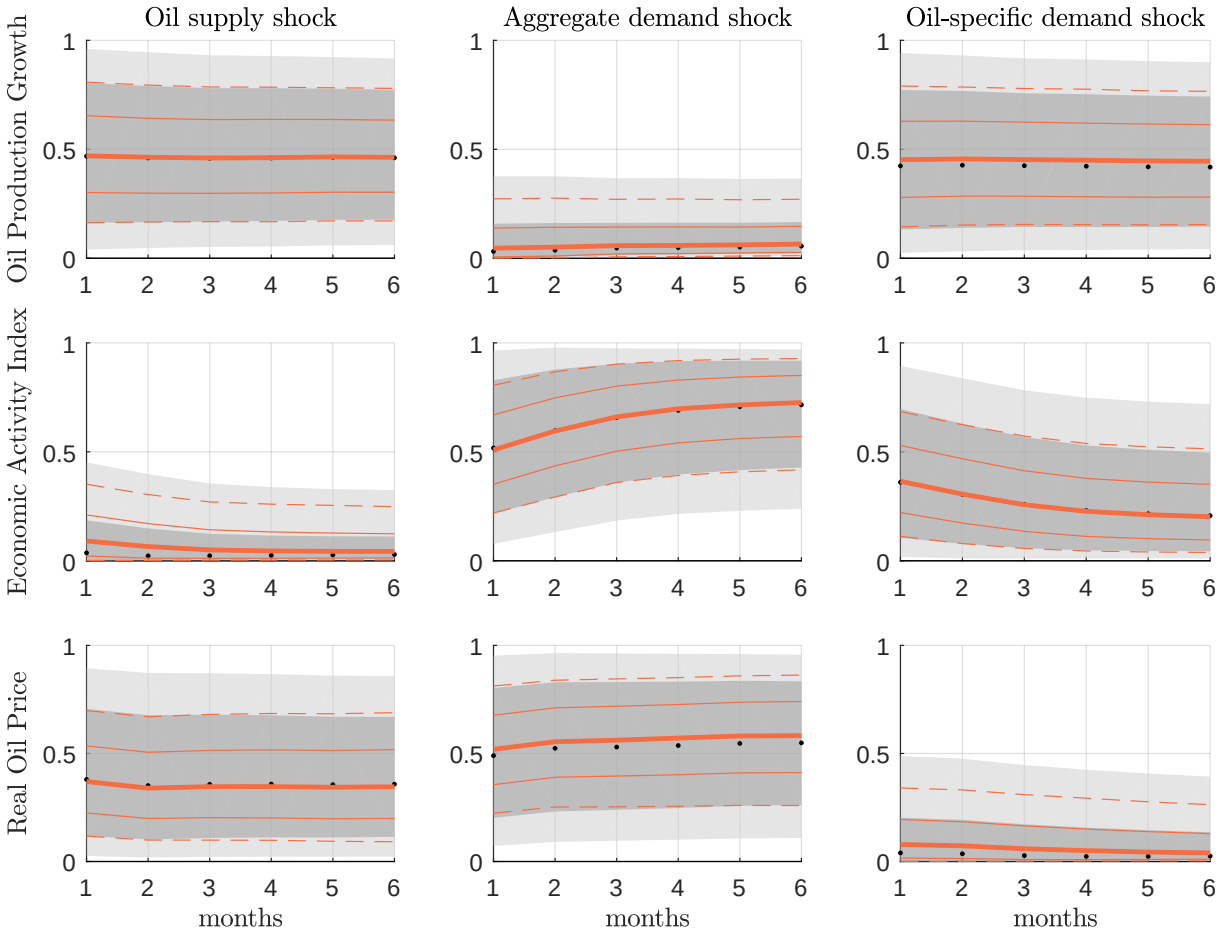
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification I.

Figure 2.39: FEVD, comparing NiWU and Np(B) (prior II)



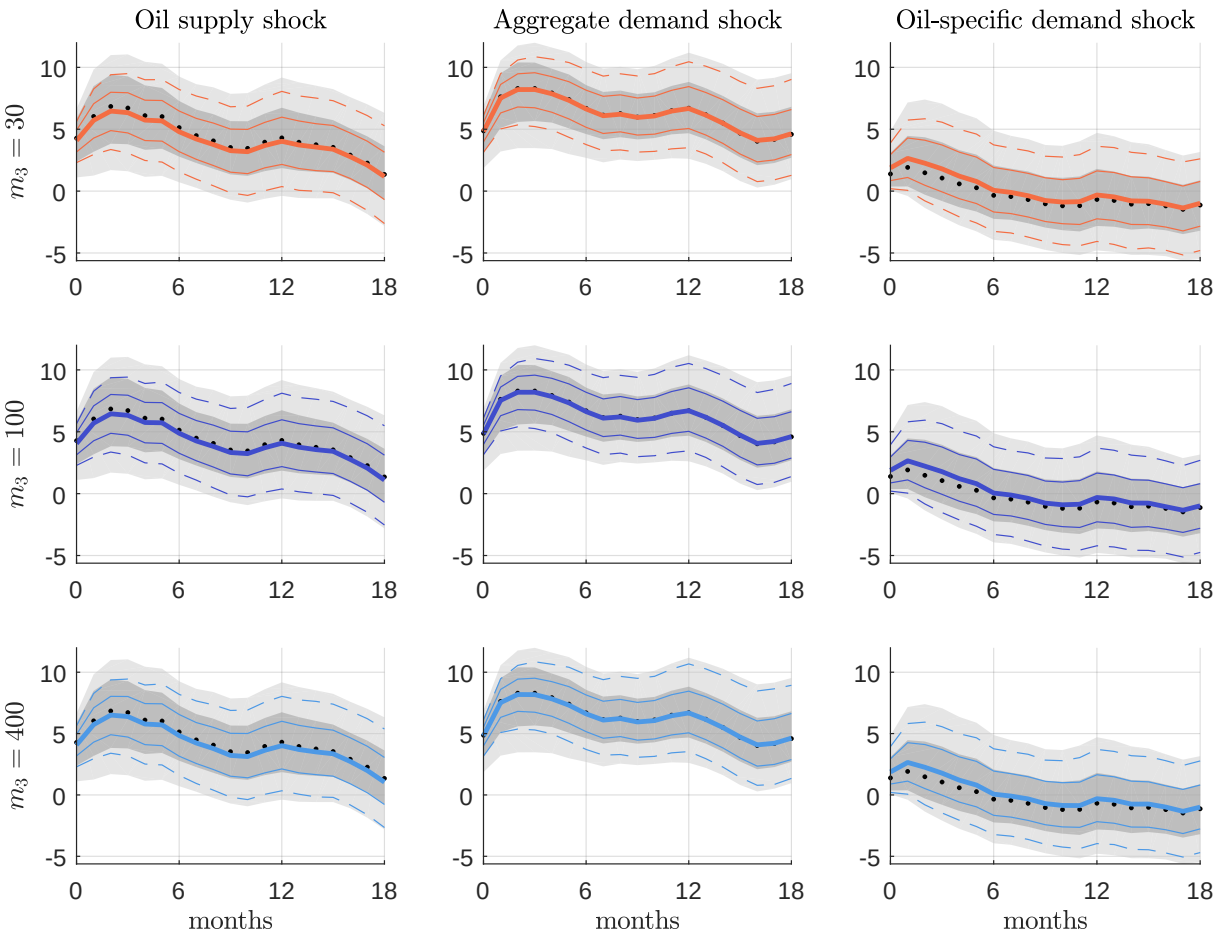
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification II.

Figure 2.40: FEVD, comparing NiWU and Np(B) (prior III)



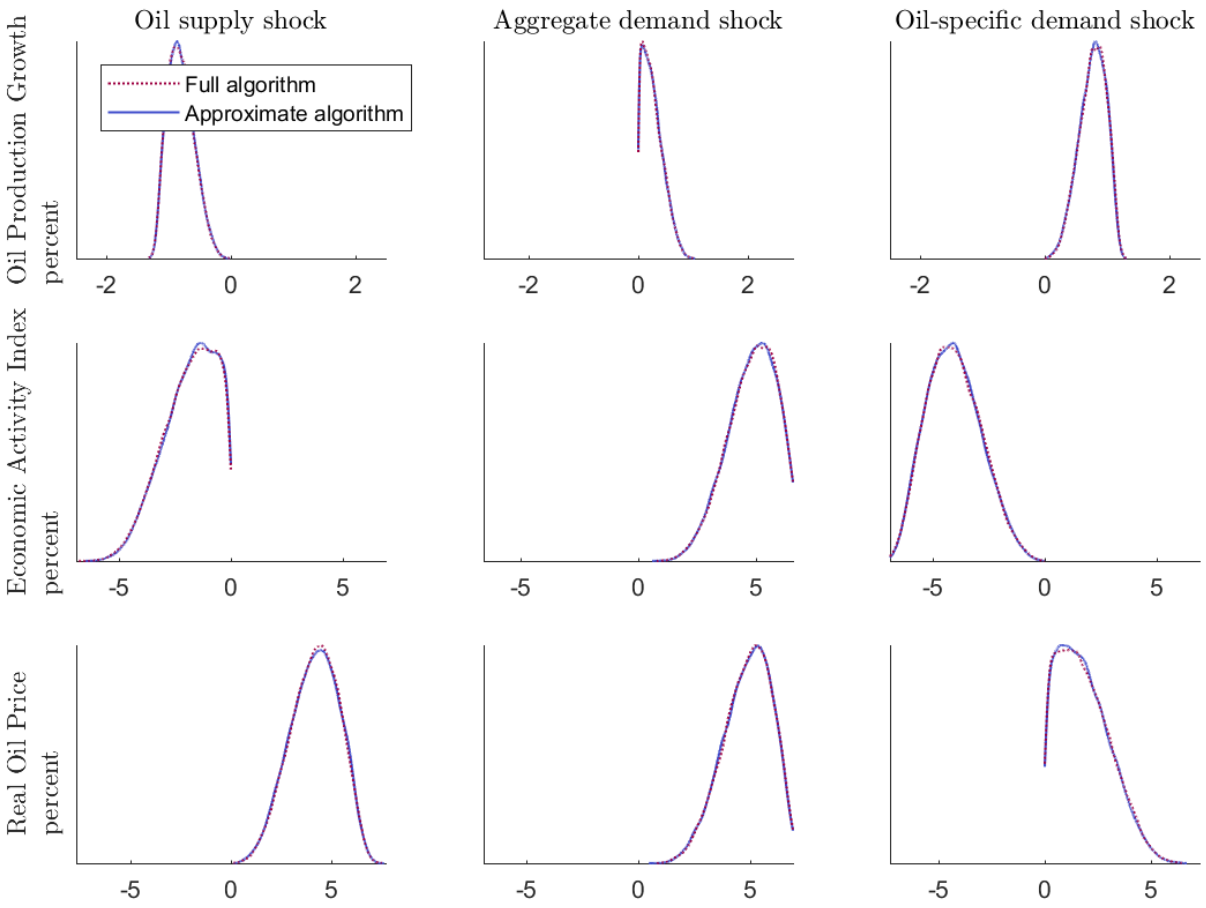
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under prior specification III.

Figure 2.41: Sensitivity analysis for figure 2.4 in the paper



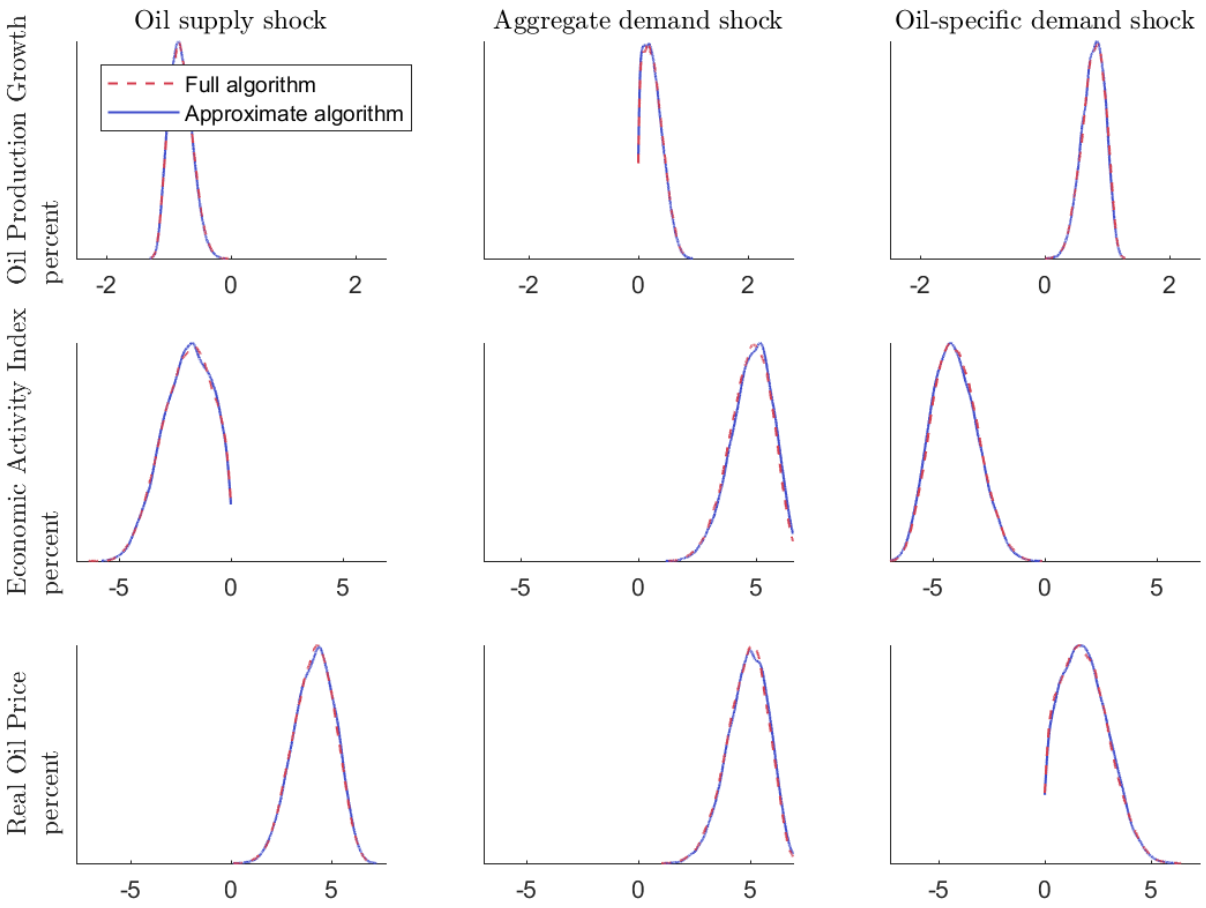
Note: The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach under Prior I. The rows of the figure differ for the number of draws used to evaluate $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB$ numerically in Stage A of our algorithm.

Figure 2.42: Sensitivity analysis. Sampling $p(B|Y)_{Np(B)}$ with our algorithm (prior I)



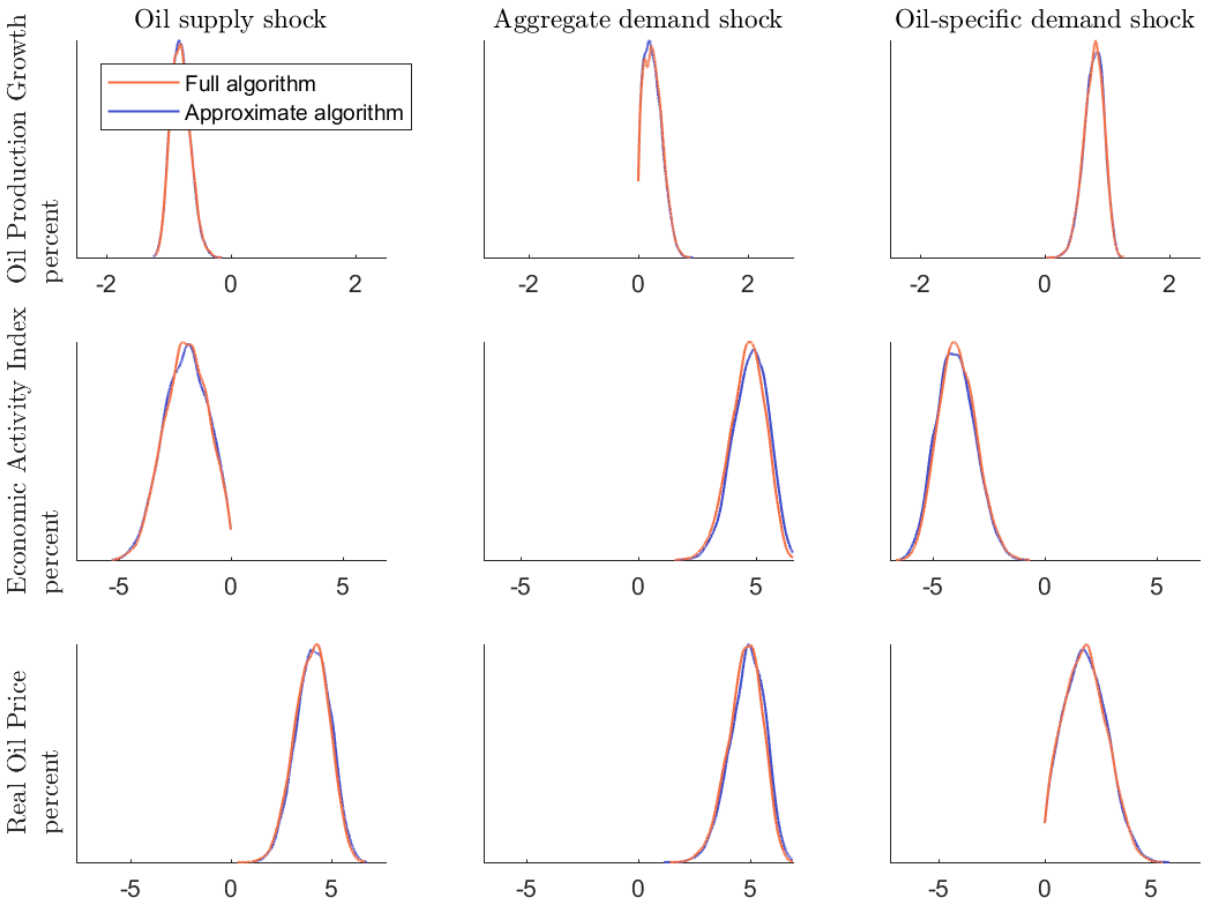
Note: The red dotted line shows the results using the full algorithm. The blue continuous line shows the results using the approximation with $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB = 1$. This approximation reduces the computational time from 34m19s to 6m6s.

Figure 2.43: Sensitivity analysis. Sampling $p(B|Y)_{Np(B)}$ with our algorithm (prior II)



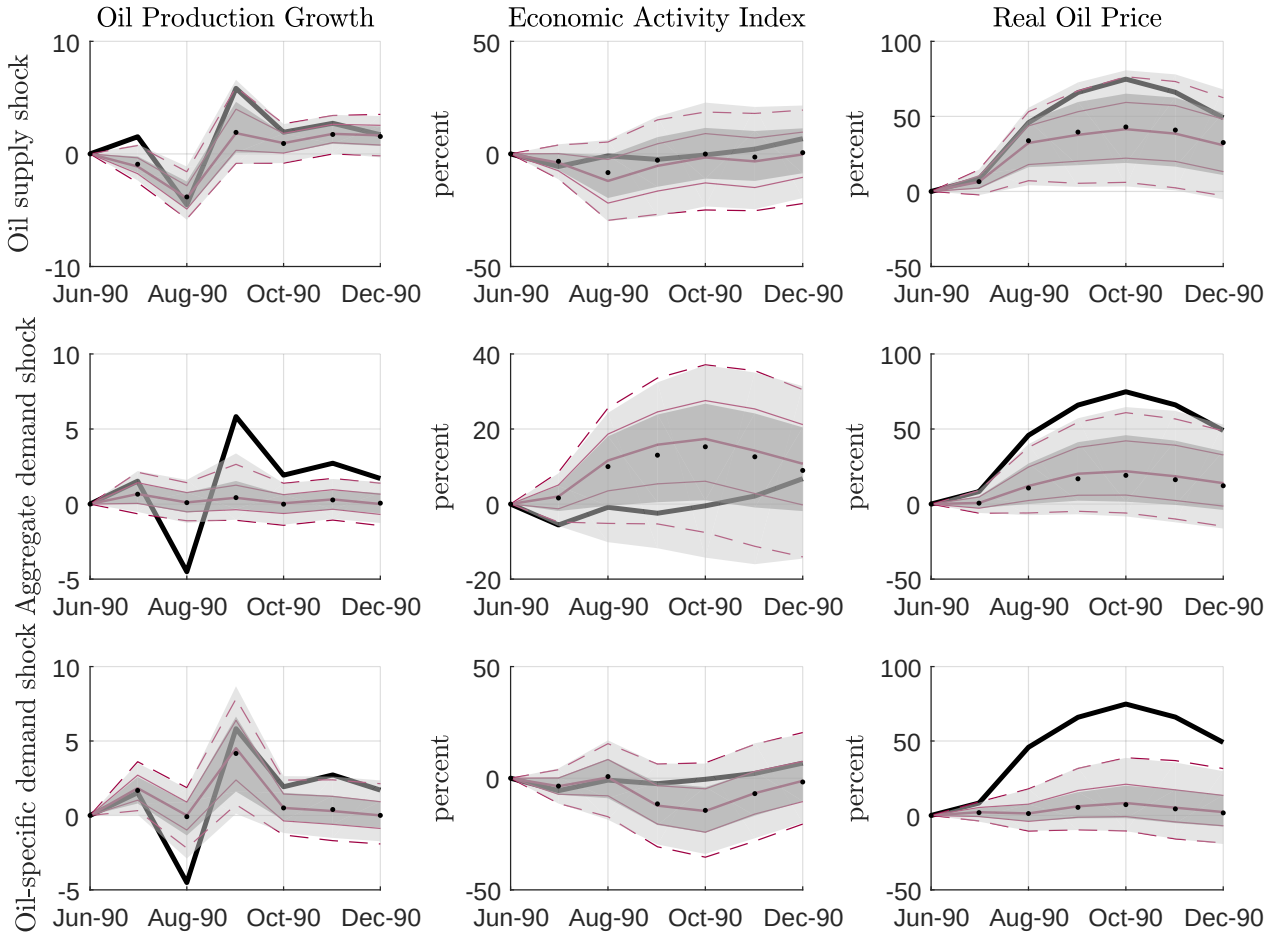
Note: The red dashed line shows the results using the full algorithm. The blue continuous line shows the results using the approximation with $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB = 1$. This approximation reduces the computational time from 35m2s to 6m5s.

Figure 2.44: Sensitivity analysis. Sampling $p(B|Y)_{Np(B)}$ with our algorithm (prior III)



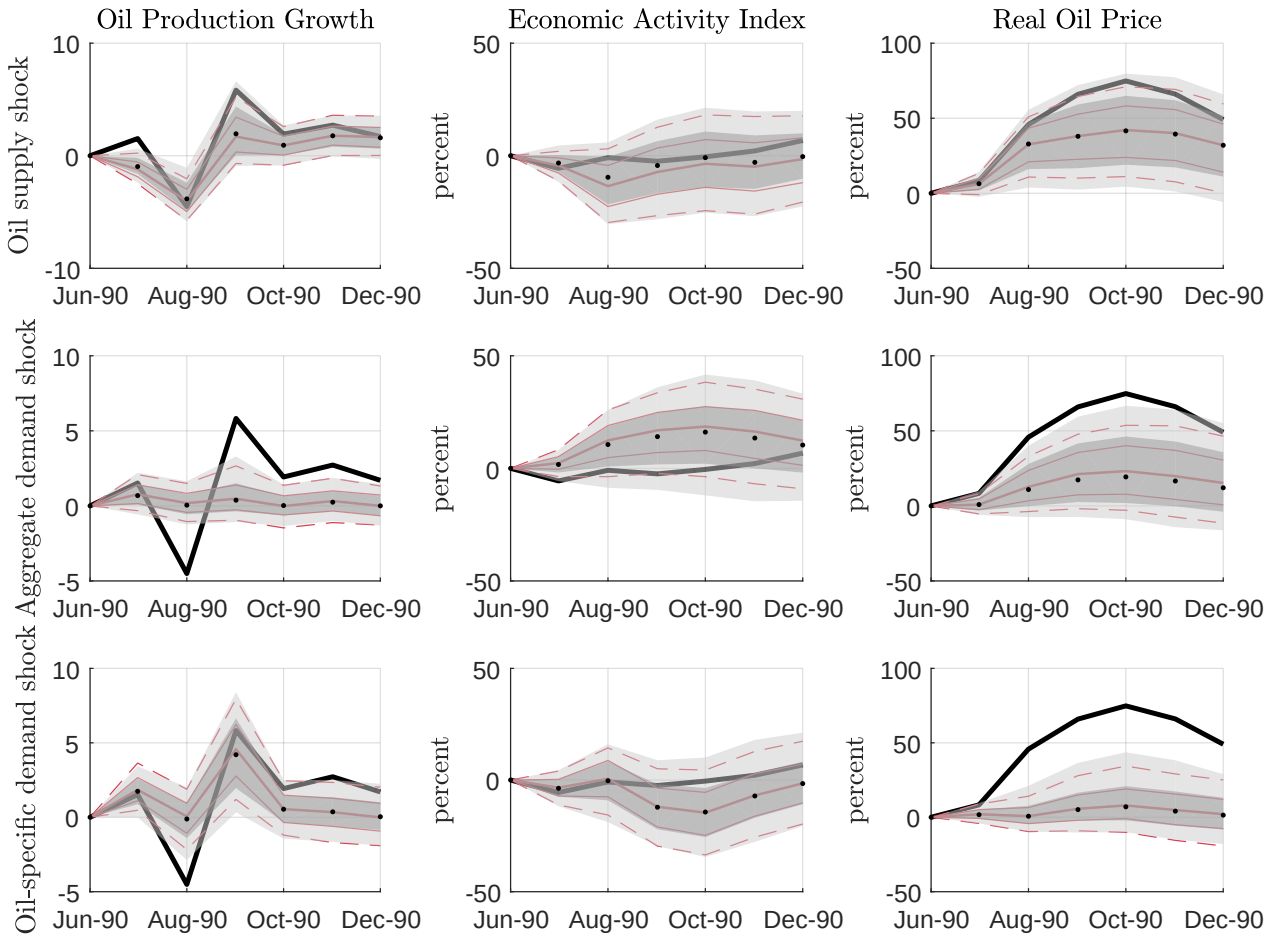
Note: The red continuous line shows the results using the full algorithm. The blue continuous line shows the results using the approximation with $\int_{\mathcal{B}(\Sigma_d)} p(B)_{Np(B)} dB = 1$. This approximation reduces the computational time from 36m21s to 8m56s.

Figure 2.45: Historical decomposition, June – December 1990 (prior I)



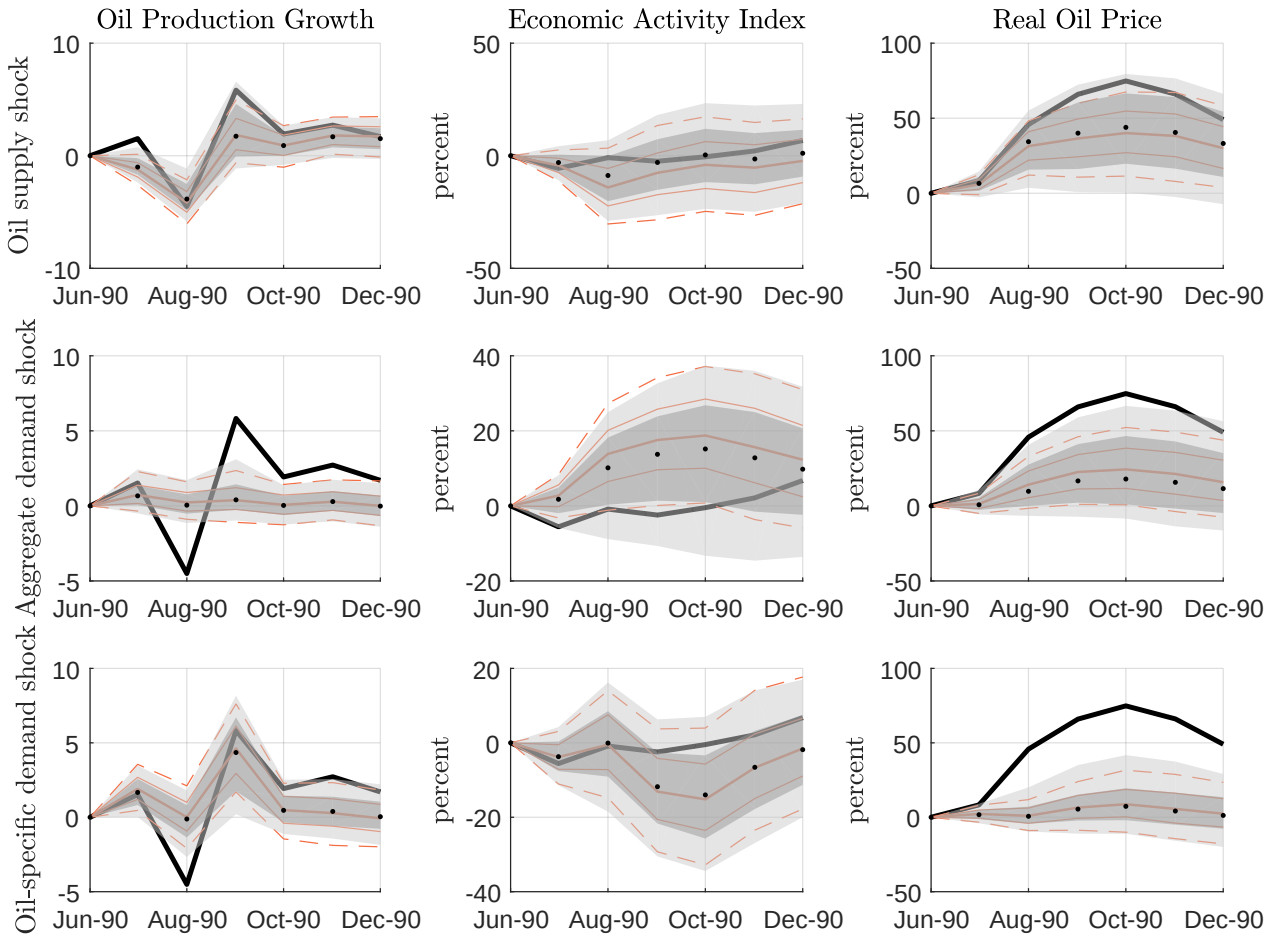
Note: The figure shows the data (solid black line) and its decomposition into the cumulative contribution of the estimated structural shocks from the beginning of the sample until period t. The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. Having subtracted the value corresponding to June 1990 before computing pointwise statistics, the figure can be interpreted as percent relative to June 1990.

Figure 2.46: Historical decomposition, June – December 1990 (prior II)



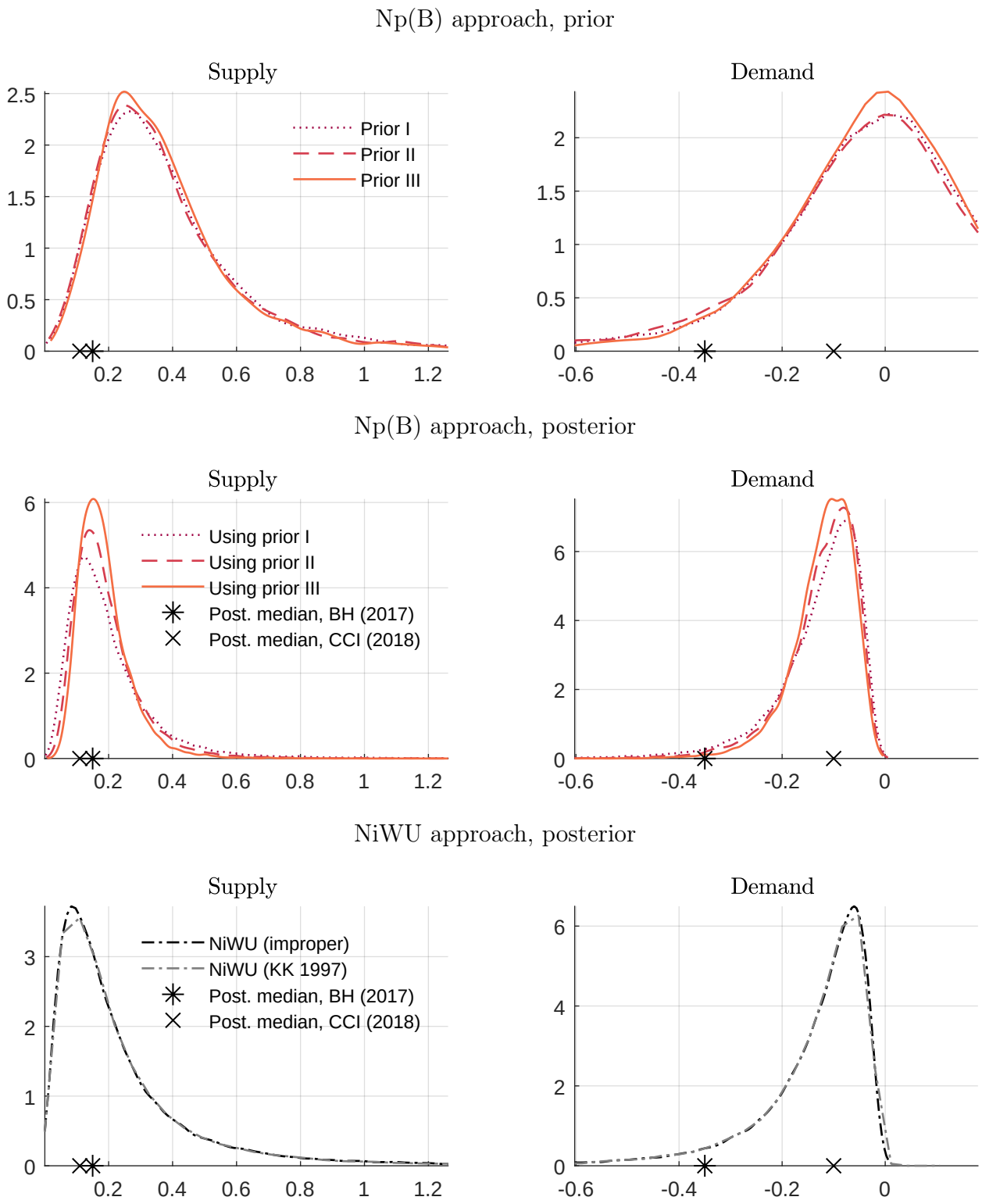
Note: The figure shows the data (solid black line) and its decomposition into the cumulative contribution of the estimated structural shocks from the beginning of the sample until period t. The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. Having subtracted the value corresponding to June 1990 before computing pointwise statistics, the figure can be interpreted as percent relative to June 1990.

Figure 2.47: Historical decomposition, June – December 1990 (prior III)



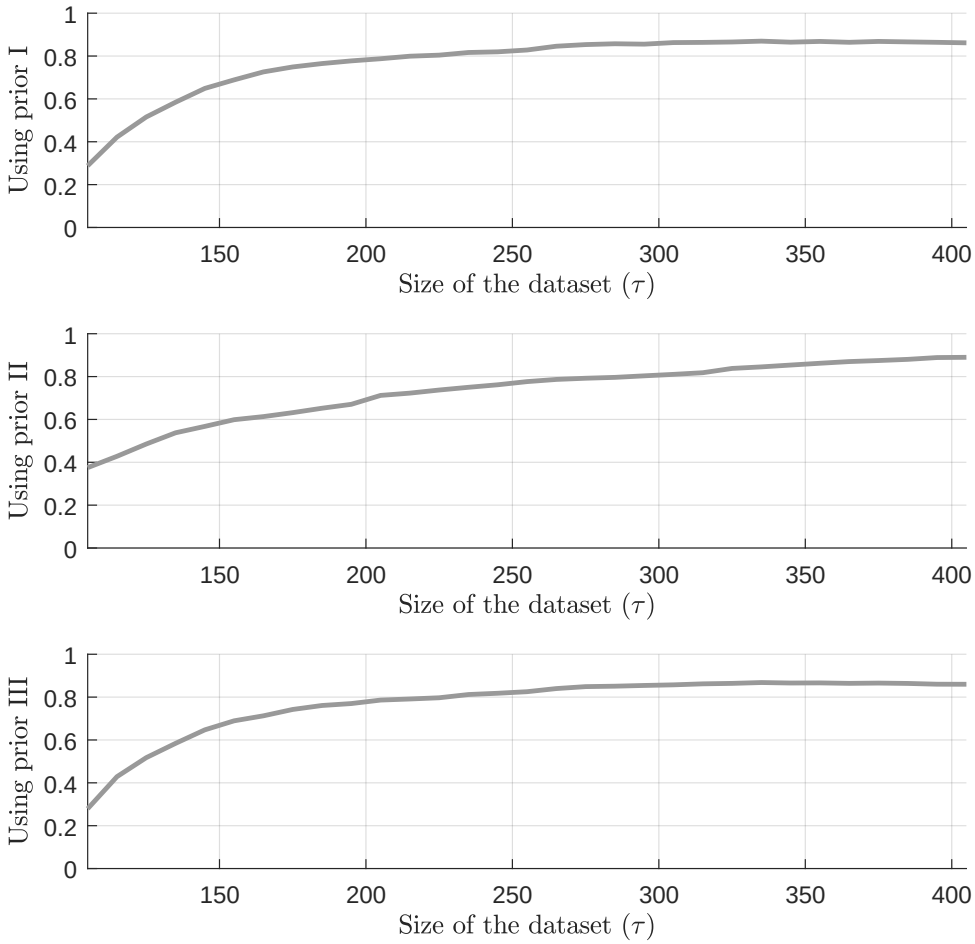
Note: The figure shows the data (solid black line) and its decomposition into the cumulative contribution of the estimated structural shocks from the beginning of the sample until period t. The dotted line and the shaded areas show the pointwise median, 68 and 95% credible bands associated with the improper prior parametrization of the NiWU approach. The remaining solid and dashed lines show the same statistics associated with our Np(B) approach. Having subtracted the value corresponding to June 1990 before computing pointwise statistics, the figure can be interpreted as percent relative to June 1990.

Figure 2.48: Price and demand elasticities



Note: The figure shows the posterior medians estimated by Baumeister and Hamilton (2019) and by Caldara et al. (2018).

Figure 2.49: Behaviour of the relevant effective sample size in Stage A of our algorithm when the size of the dataset increases



Note: The figure reports the ratio of the effective sample size relative to the proposal draws from Stage A of our algorithm. The training sample used to set the prior beliefs is the same as in the application in the paper. We then use the first $p + \tau$ observations of the estimation sample and progressively increase τ, starting from $\tau = 105$. The ratio reaches approximately 0.50 for $\tau = 125$ for all three cases. The Gibbs sampler extracts 20,000 burn-in draws and 40,000 retained draws to reduce computational time.

Table 2.14: Performance of our algorithm

	Stage A			Stage B		
	proposal draws	effective sample size	relative effective sample size	proposal draws	effective sample size	relative effective sample size
	m_2	ESS^A	$\frac{ESS^A}{m_2}$	m_5	ESS^B	$\frac{ESS^B}{m_5}$
Prior I	50,000	42,947	0.8589	80,000	59,314	0.7414
Prior II	50,000	44,618	0.8924	80,000	37,698	0.4712
Prior III	50,000	34,286	0.6857	80,000	15,971	0.1996

Table 2.15: Comparison of the computational time

	Np(B) approach												**NiWU approach**					
	Our algorithm									DSMH algorithm			with improper prior			with KK(1997) prior		
	Stage A			Stage B			Total											
	h	m	s	h	m	s	h	m	s	h	m	s	h	m	s	h	m	s
Prior I		32	44		1	34		34	19	24	52	29		6	11		6	11
Prior II		33	23		1	39		35	2	25	29	42		6	11		6	11
Prior III		34	48		1	33		36	21	25	54	11		6	11		6	11

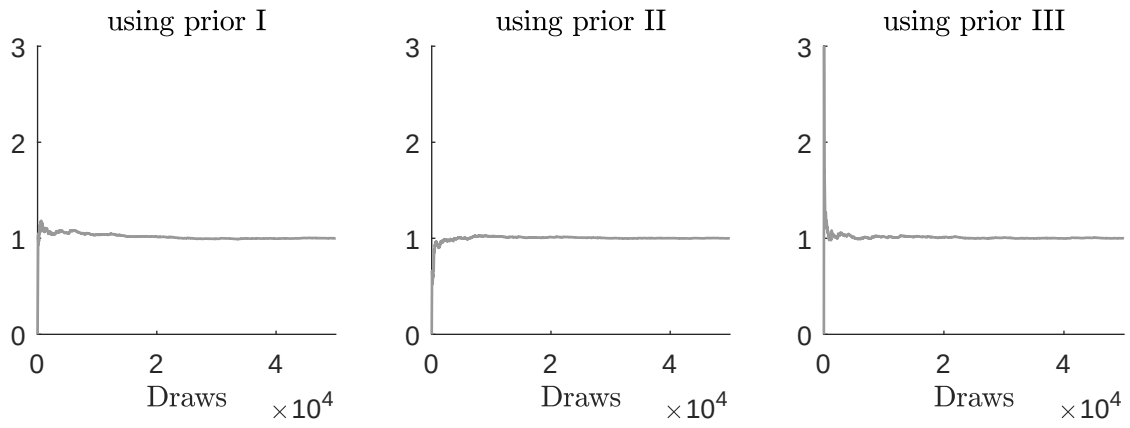
Note: All codes are run on Matlab, except for the Dynamic Striated Metropolis-Hastings algorithm, which we coded on Fortran to reduce computational time.

Table 2.16: Diagnostics on the importance weights, tests

	u	0.5N	0.6N	0.7N	0.9N	0.99N
using prior I						
	Wald	-51.44	-47.03	-41.35	-23.86	-7.09
	Score	-21.52	-18.49	-15.25	-8.05	-2.50
	LR	0.00	0.00	0.00	0.00	0.00
using prior II						
	Wald	-53.55	-49.36	-43.88	-26.37	-8.25
	Score	-22.03	-18.98	-15.81	-8.61	-2.45
	LR	0.00	0.00	0.00	0.00	0.00
using prior III						
	Wald	-48.25	-43.28	-37.43	-20.99	-5.93
	Score	-19.49	-16.95	-14.36	-7.91	-2.54
	LR	0.00	0.00	0.00	0.00	0.00

Note: Reported are the test statistics. The null hypothesis implies finite weight variance. The corresponding critical values above which the null hypothesis is rejected are 1.65 for the Wald test, 1.65 for the score test and 2.68 for the LR test (5% significance level). The corresponding p-values are close to 1 in all cases.

Figure 2.50: Diagnostics on the importance weights, graphical assessment



Note: The graph shows the recursive variance $\{v_i\}_{i=1}^{N}$, where $v_i = var(w_{1:i})$ computed using de-meaned and standardized weights.

Table 2.17: Convergence for the application in section 2.4 of the paper (prior I)

parameter	1	2	3	4	5	6	7	8	9	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01										
stage 1	0.89	0.99	0.76	0.52	0.97	0.81	0.89	0.93	0.60	
stage 2	0.34	0.10	0.91	0.70	0.45	0.53	0.50	0.95	0.32	
stage 3	0.97	0.00	0.44	0.38	0.73	0.11	0.20	0.84	0.77	
stage 4	0.34	0.13	0.07	0.10	0.02	0.00	0.53	0.11	0.19	
stage 5	0.04	0.47	0.30	0.80	0.43	0.37	0.09	0.20	0.08	
stage 6	0.51	0.22	0.73	0.89	0.13	0.84	0.13	0.48	0.08	
stage 7	0.24	0.09	0.91	0.42	0.54	0.88	0.13	0.10	0.67	
stage 8	0.71	0.87	0.35	0.67	0.53	0.84	0.82	0.83	0.22	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (96,000)										
stage 1	41,344	46,534	49,785	35,215	51,856	51,204	54,525	38,755	35,627	
stage 2	37,434	50,066	53,377	36,224	54,932	46,233	56,568	36,924	37,273	
stage 3	38,730	44,984	56,174	34,828	50,566	53,771	57,904	39,664	35,354	
stage 4	34,979	48,661	47,399	31,605	57,975	46,849	50,863	39,008	34,591	
stage 5	33,342	54,051	48,941	29,899	45,119	48,833	48,579	36,675	31,792	
stage 6	32,761	44,036	44,523	28,301	48,893	51,204	46,194	34,675	31,982	
stage 7	32,849	42,128	50,356	28,183	46,572	47,749	49,348	36,461	29,521	
stage 8	31,937	42,577	47,819	28,189	46,003	40,744	44,222	32,982	27,444	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$										
stage 1	1.001	1.002	1.001	1.001	1.001	1.001	1.001	1.001	1.002	
stage 2	1.001	1.001	1.001	1.000	1.001	1.001	1.001	1.001	1.001	
stage 3	1.001	1.003	1.001	1.001	1.001	1.001	1.002	1.001	1.001	
stage 4	1.001	1.004	1.002	1.003	1.002	1.002	1.002	1.002	1.001	
stage 5	1.002	1.002	1.001	1.002	1.001	1.001	1.001	1.001	1.003	
stage 6	1.001	1.001	1.000	1.001	1.001	1.000	1.001	1.001	1.003	
stage 7	1.000	1.002	1.001	1.004	1.001	1.001	1.002	1.002	1.001	
stage 8	1.006	1.001	1.005	1.002	1.002	1.002	1.004	1.002	1.003	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$										
stage 1										1.005
stage 2										1.004
stage 3										1.006
stage 4										1.006
stage 5										1.005
stage 6										1.005
stage 7										1.006
stage 8										1.009

Table 2.18: Convergence diagnostics for the application in section 2.4 of the paper (prior II)

parameter	1	2	3	4	5	6	7	8	9	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01										
stage 1	0.89	0.99	0.91	0.77	0.76	0.63	0.99	0.97	0.22	
stage 2	0.43	0.83	0.69	0.51	0.74	0.38	0.64	0.52	0.92	
stage 3	0.43	0.22	0.96	0.16	0.72	0.64	0.30	0.19	0.97	
stage 4	0.83	0.46	0.47	0.10	0.39	0.12	0.60	0.47	0.87	
stage 5	0.00	0.06	0.04	0.23	0.87	0.40	0.00	0.00	0.09	
stage 6	0.47	0.92	0.92	0.74	0.80	0.72	0.39	0.92	0.66	
stage 7	0.40	0.03	0.18	0.28	0.89	0.20	0.11	0.32	0.43	
stage 8	0.01	0.89	0.09	0.78	0.01	0.03	0.07	0.02	0.33	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (96,000)										
stage 1	38,119	47,370	45,685	33,838	48,710	54,814	50,787	38,627	35,620	
stage 2	36,111	51,075	45,272	36,428	53,673	52,694	50,161	38,833	36,033	
stage 3	32,480	52,883	47,749	32,248	47,931	47,439	45,531	36,582	34,444	
stage 4	34,444	45,531	44,619	31,061	46,337	47,739	46,519	36,616	29,441	
stage 5	32,019	43,502	46,992	30,289	47,548	44,677	46,233	34,586	31,792	
stage 6	35,469	43,375	41,654	29,404	40,952	47,123	50,108	32,027	29,643	
stage 7	31,382	41,270	43,832	30,067	41,937	45,093	47,313	32,065	28,585	
stage 8	30,502	38,932	43,630	31,056	44,353	44,017	40,908	34,958	28,010	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$										
stage 1	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.002	1.001	
stage 2	1.002	1.000	1.002	1.001	1.001	1.001	1.001	1.001	1.001	
stage 3	1.002	1.003	1.001	1.003	1.002	1.003	1.002	1.004	1.001	
stage 4	1.001	1.001	1.001	1.002	1.001	1.001	1.001	1.002	1.001	
stage 5	1.003	1.003	1.004	1.004	1.001	1.003	1.003	1.002	1.003	
stage 6	1.003	1.001	1.002	1.002	1.002	1.001	1.001	1.001	1.003	
stage 7	1.002	1.002	1.002	1.001	1.001	1.003	1.002	1.001	1.002	
stage 8	1.001	1.002	1.001	1.000	1.002	1.002	1.001	1.002	1.001	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$										
stage 1										1.004
stage 2										1.005
stage 3										1.006
stage 4										1.006
stage 5										1.007
stage 6										1.006
stage 7										1.007
stage 8										1.004

Table 2.19: Convergence diagnostics for the application in section 2.4 of the paper (prior III)

parameter	1	2	3	4	5	6	7	8	9	all
Geweke (1992), p-value. Convergence found if p-value ≥ 0.01										
stage 1	0.34	0.19	0.79	0.21	0.15	0.93	0.45	0.69	0.23	
stage 2	0.19	0.05	0.77	0.08	0.31	0.40	0.16	0.02	0.26	
stage 3	0.02	0.66	0.34	0.85	0.30	0.05	0.11	0.01	0.57	
stage 4	0.83	0.21	0.22	0.13	0.91	0.66	0.37	0.39	0.70	
stage 5	0.58	0.65	0.63	0.11	0.13	0.35	0.75	0.13	0.28	
stage 6	0.11	0.00	0.56	0.99	0.26	0.52	0.03	0.33	0.10	
stage 7	0.96	0.54	0.83	0.31	0.97	0.98	0.74	0.72	0.74	
stage 8	0.84	0.94	0.75	0.73	0.29	0.39	0.99	0.39	0.91	
Raftery and Lewis (1992), n^*. Convergence found if $n^* \leq N \cdot G$ (96,000)										
stage 1	37,689	53,538	51,722	34,207	44,543	54,581	49,135	42,587	37,145	
stage 2	34,567	45,578	46,664	33,528	55,295	49,189	47,142	41,561	37,241	
stage 3	33,222	43,924	48,205	32,954	45,218	46,233	51,032	33,876	30,648	
stage 4	31,323	43,539	46,376	29,732	43,055	47,250	47,880	33,629	27,675	
stage 5	31,199	41,190	42,998	30,821	47,182	42,177	46,664	39,145	27,893	
stage 6	31,334	37,051	46,625	27,926	49,575	43,758	46,625	35,121	27,064	
stage 7	28,247	35,268	37,146	26,598	37,434	38,522	41,190	30,424	26,591	
stage 8	28,836	34,149	33,992	26,167	32,697	35,645	35,472	29,141	27,444	
Gelman and Rubin (1992), $\hat{R}$. Convergence found if $\hat{R} \leq 1.2$										
stage 1	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.002	1.000	
stage 2	1.002	1.002	1.001	1.003	1.003	1.002	1.003	1.005	1.002	
stage 3	1.002	1.002	1.001	1.002	1.000	1.001	1.002	1.001	1.001	
stage 4	1.001	1.002	1.001	1.001	1.002	1.001	1.002	1.002	1.001	
stage 5	1.001	1.001	1.000	1.001	1.001	1.001	1.001	1.001	1.001	
stage 6	1.002	1.002	1.001	1.001	1.000	1.001	1.002	1.001	1.001	
stage 7	1.002	1.001	1.002	1.001	1.000	1.001	1.001	1.001	1.001	
stage 8	1.001	1.001	1.001	1.002	1.001	1.000	1.001	1.001	1.002	
Brooks and Gelman (1998), $\hat{R}^{mult}$. Convergence found if $\hat{R}^{mult} \leq 1.2$										
stage 1										1.004
stage 2										1.006
stage 3										1.005
stage 4										1.006
stage 5										1.004
stage 6										1.003
stage 7										1.005
stage 8										1.005

CHAPTER 3

Combining Factor Models and External Instruments to Identify Uncertainty Shocks

3.1 Introduction

Following the seminal paper by Bloom (2009) a fast growing literature analyses the macroeconomic impact of exogenous increases in uncertainty using structural VAR models. An increase in uncertainty is broadly defined as increased difficulties of economic agents to make accurate forecasts. Within this literature, there is a consensus that exogenous increases in uncertainty lead to adverse real effects. These include falling production, hours, and employment. However, there is an ongoing debate about whether these reactions are dominated by supply or demand channels, i.e. whether they are accompanied by a rise or fall in inflation. These nominal reactions are crucial for policy makers: if, for example, central banks know that a rise in uncertainty leads to a decrease in inflation, they can, theoretically, move both real and nominal variables back to their desired targets by employing an expansionary policy. If, on the other hand, uncertainty shocks do not affect prices or are inflationary, central banks are faced with a trade-off between allowing for more inflation or a decline in real activity.

In this paper, I propose a Bayesian Proxy Factor-augmented VAR (BP-FAVAR) model to analyse the real and nominal effects following an uncertainty shock. This novel model offers a unified framework to combine a large information set with a non-recursive identification strategy. It addresses two shortcomings in commonly used small-scale, recursively identified, VAR models: (i) informational insufficiency and (ii) non-credible identification. I find that inflation responds negatively to a positive uncertainty shock in the short run and is indistinguishable from zero after six months. This is comforting news for policy makers given that they can address both real and nominal effects using standard instruments. The dynamic effects depend strongly on the identification

scheme. Biases resulting from a recursive scheme cannot be alleviated by augmenting the information set of the model.

The structural VAR literature is inconclusive about the inflationary effects of uncertainty shocks. Leduc and Liu (2016), using a four-variable, recursively identified VAR model, find that uncertainty shocks are deflationary, even in the medium term. Piffer and Podstawski (2017), identifying the work-horse model by Bloom (2009) via an external instrument, find a short-lived drop and fast rebound in prices. Caggiano, Castelnuovo and Nodari (2017), extending a small-scale VAR model to a non-linear setting, find that uncertainty shocks are deflationary only in recessions and have no effect on prices in expansions. Caggiano, Castelnuovo and Pellegrino (2017), employing an interacted VAR model, find that the price reaction is indistinguishable from zero over the whole impulse response horizon. Carriero et al. (2018) distinguish between macroeconomic and financial uncertainty and find that while increases in the former are inflationary, increases in the latter tend to be deflationary.

At best, the theoretical literature provides limited guidance for the inflationary effects of uncertainty shocks. Fernández-Villaverde et al. (2015) and Born and Pfeifer (2014) put forward two opposing channels to explain the potential price reaction following an exogenous increase in policy uncertainty: On the one hand, if consumers face difficulties predicting the next period, they will postpone consumption decisions, which will lead to a fall in both economic activity and prices. Therefore, an uncertainty shock would resemble an aggregate demand shock. Using a model with labour market frictions, Leduc and Liu (2016) also reach this conclusion. On the other hand, if firms face difficulties predicting the next period, they will bias their price decision upwards. The reason is that their profit function is concave in prices, making it more costly to set prices too low rather than too high. If this "price-bias-channel" dominates, the reactions of real and nominal variables will resemble a short-lived aggregate supply shock. The aggregate reaction of prices from these two opposing channels is ambiguous.

The responses to an uncertainty shock may depend heavily on the information set of the model, as pointed out by Caggiano, Castelnuovo and Nodari (2017) and Angelini et al. (2019). In particular, omitted variables, such as consumer sentiment (Sims, 2012), total factor productivity (Bachmann and Bayer, 2013), and measures of anticipated risk (Christiano et al., 2014) may bias the impulse responses. In order to avoid having to add potentially omitted variables one by one, I augment the work-horse VAR model by Bloom (2009) with latent factors. These summarise the information contained in a large set of variables and, thus, should alleviate omitted variable biases. Second, as pointed out by Stock and Watson (2012), a recursive identification scheme may be invalid given the contemporaneous interrelation between the real economy and uncertainty

as well as the fast moving nature of financial markets. Therefore, a recent strand of the literature addresses identification issues by departing from recursive schemes and employing external instruments. I follow Piffer and Podstawski (2017) in identifying an uncertainty shock using a proxy based on the price of gold. This proxy captures movements in the price of gold around selected economic and political events. These events are associated with movements in uncertainty. Given that gold can be considered a safe haven asset, these movements should capture exogenous variations in uncertainty.

Estimation of the model is subject to two challenges. The first regards the so-called "curse of dimensionality". Even after shrinking the variable space using latent factors, the model still contains a large number of parameters. The baseline model consists of over 800 parameters, while the effective sample length is only roughly 400. Therefore, estimation is challenging in a frequentist setting. A second challenge arises from the need to effectively summarise the estimation uncertainty in both the model parameters and the latent factors. This is difficult using bootstrap techniques (see for example Yamamoto, 2019). In addition, there are no asymptotic results justifying the use of such techniques, as pointed out by Kilian and Lütkepohl (2017). To jointly address these two challenges I employ a Bayesian approach. It allows for overcoming dimensionality problems by shrinking the parameter space and summarises the estimation uncertainty in a joint posterior distribution. The BP-FAVAR can be considered a combination of the Bayesian FAVAR estimation proposed by Belviso and Milani (2006) and the Bayesian Proxy VAR by Caldara and Herbst (2019). I re-parametrize their model to impose structure on the impact effects of shocks.

The main results are the following: Uncertainty shocks are deflationary in the short run. The price reaction is indistinguishable from zero after about six months. Real variables and the stock market drop and rebound. This suggests that policy makers can employ an expansionary policy to alleviate the adverse effects of an exogenous increase of uncertainty on both prices and real activity. I show evidence that the workhorse model by Bloom (2009) is informationally deficient. When computing impulse responses, I find that alleviating informational deficiency problems has only marginal quantitative effects. This finding is in line with Sims (2012), who points out that informational deficiency is not an either/or but a quantitative issue. In the present case the biases it causes are negligible, which is comforting news.

The remainder of the paper is organized as follows: Section 2 introduces the model setup and explains the identification of uncertainty shocks. Section 3 presents the data and discusses the results. The last section concludes.

3.2 The Bayesian Proxy FAVAR

In this section, I introduce the Bayesian Proxy FAVAR model. I start by describing the different parts of the model. Then, I discuss how partial identification is achieved. Lastly, I show how the model can be decomposed into three blocks to facilitate inference.

3.2.1 Model Description

The Bayesian Proxy Factor-augmented VAR model admits a state-space form, which consists of an observation equation, a transition equation and a proxy equation. First, consider the observation equation, which shows how latent and observable factors map into informational series:

$$\boldsymbol{x}_t = \Lambda^f \boldsymbol{f}_t + \Lambda^z \boldsymbol{z}_t + \boldsymbol{\xi}_t \tag{3.1}$$

$$\boldsymbol{\xi}_t \sim N(\boldsymbol{0}, \Omega) \tag{3.2}$$

where $\boldsymbol{x}_t$ is a $N \times 1$ vector of observable series, $\boldsymbol{f}_t$ is a $R \times 1$ vector of latent factors, and $\boldsymbol{z}_t$ is a $K \times 1$ vector of observable factors. Importantly, $\boldsymbol{x}_t$ does not contain any of the observable factors in $\boldsymbol{z}_t$. Λ^f is a $N \times R$ matrix of factor loadings for latent factors and Λ^z is a $N \times K$ matrix of coefficients for the observable factors. $\boldsymbol{\xi}_t$ is a $N \times 1$ vector of idiosyncratic errors. In general, $\boldsymbol{\xi}_t$ can be serially correlated, i.e. $Cov(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-j}) \neq 0$ for some j, but they are uncorrelated across series, i.e. $Var(\boldsymbol{\xi}_t) = \Omega$ is assumed to be diagonal.

Next, consider the transition equation which shows the dynamic evolution of the factors. It writes as a VAR(P) of the following form:

$$\boldsymbol{y}_t = \Pi \boldsymbol{w}_t + \boldsymbol{u}_t \tag{3.3}$$

$$\boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma), \tag{3.4}$$

where $\boldsymbol{y}_t = \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \end{bmatrix}$ stacks latent and observable factors in a vector. The coefficient matrix $\Pi = [c, \Pi_1, ..., \Pi_P]$ of dimension $(R+K) \times (P(R+K)+1)$ contains the autoregressive parameters of the VAR. $\boldsymbol{w}_t = [\boldsymbol{1}_{R+K\times 1}; \boldsymbol{y}'_{t-1}, ..., \boldsymbol{y}'_{t-P}]'$ stacks a constant and P lags of $\boldsymbol{y}_t$. The $(R+K) \times 1$ vector of reduced form errors, $\boldsymbol{u}_t$, is serially uncorrelated, i.e. $Cov(\boldsymbol{u}_t, \boldsymbol{u}_{t-p}) = 0 \quad \forall t = 1, ..., T, \forall p = 1, ..., \infty$. Also, $\boldsymbol{u}_t$ are uncorrelated with all leads and lags of the idiosyncratic errors, $\boldsymbol{\xi}_t$, i.e. $Cov(\boldsymbol{u}_t \boldsymbol{\xi}_{t-j}) = 0 \; \forall j = 1, ..., \infty$, $\forall t = 1, ..., T$.

I impose structure on the on-impact effects of structural shocks by assuming that the reduced form errors map into structural shocks as:

$$\boldsymbol{u}_t = B\boldsymbol{\epsilon}_t \tag{3.5}$$

$$\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, I_{R+K}), \tag{3.6}$$

where B is a $(R+K) \times (R+K)$ matrix containing the on-impact effects of the structural shocks. Their variance is normalised to one and they are contemporaneously uncorrelated. This implies the following relation between the reduced form covariance matrix and the matrix of on-impact effects: $\Sigma = BB'$

As is well known, further restrictions beyond those implied by the covariance matrix are needed to identify B. The reason is that the data cannot discriminate between observationally equivalent representations: All B such that $BB' = \Sigma$ yield the same likelihood.

In order to identify the first column of B, which I denote by $\boldsymbol{b}$, I augment the model by a "proxy equation", as in Caldara and Herbst (2019). It spells out the relation between structural shock and instrument and is given as[1]:

$$m_t = \beta\epsilon_{1,t} + \sigma_\nu \nu_t \tag{3.7}$$

$$\nu_t \sim N(0, 1), \tag{3.8}$$

where m_t is a scalar instrument correlated with the shock of interest, $\epsilon_{1,t}$. The shock of interest is ordered first, without loss of generality. Furthermore, m_t is orthogonal to all other shocks, $\boldsymbol{\epsilon}_{-1,t}$, i.e. $E(m_t\boldsymbol{\epsilon}_{-1,t}) = 0 \ \forall t$, where $\boldsymbol{\epsilon}_{-1,t}$ stands for a vector containing all but the first shock. In other words, the instrument needs to be both relevant and exogenous in order to be appropriate for identification. β captures the relationship between instrument and shock, while ν_t captures any noise contained in the instrument. The higher its variance, σ_ν^2, the less information the instrument contains about the shock of interest.

[1]Unlike their case, however, identification focuses on the on-impact effects of the shocks rather than on the contemporaneous relations of the variables included in the model. Put differently, the model imposes structure on B, rather than on B^{-1}. Caldara and Herbst (2019) estimate a so-called A-model (see Kilian and Lütkepohl, 2017 for a discussion). The A-model specification is appropriate given their aim of identifying a monetary policy equation. In the context of uncertainty shocks, however, it is more common to inform the on-impact effects of shocks (see e.g. Bloom, 2009 or Caggiano, Castelnuovo and Nodari, 2017). Therefore, I propose to use a so-called B-model, which imposes structure on the on-impact effects.

The full model can be written in compact matrix notation as:

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{z}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \Lambda^f & \Lambda^z & \mathbf{0}_{N\times 1} \\ 0_{K\times R} & I_K & \mathbf{0}_{K\times 1} \\ \mathbf{0}_{1\times R} & \mathbf{0}_{1\times K} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \\ m_t \end{bmatrix} + \begin{bmatrix} \boldsymbol{\xi}_t \\ \mathbf{0}_{K\times 1} \\ 0 \end{bmatrix} \tag{3.9}$$

$$Var(\boldsymbol{\xi}_t) = \Omega \tag{3.10}$$

$$\begin{bmatrix} \boldsymbol{y}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \Pi \\ \mathbf{0}_{1\times P(K+R)} \end{bmatrix} \boldsymbol{w}_t + \mathcal{B} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix} \tag{3.11}$$

$$Var(\begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix}) = \begin{bmatrix} I_{R+K} & \mathbf{0}_{R+K\times 1} \\ \mathbf{0}_{1\times R+K} & 1 \end{bmatrix}, \tag{3.12}$$

where $\mathcal{B} = \begin{bmatrix} B & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & \sigma_\nu \end{bmatrix}$, and $\boldsymbol{\beta} = \begin{bmatrix} \beta \\ \mathbf{0}_{(R+K-1)\times 1} \end{bmatrix}$.

3.2.2 Identification

Shock identification in the BP-FAVAR model is achieved by weighting draws from the posterior of structural parameters. In particular, more weight is given to posterior draws which lead to a close relation between instrument and the shock of interest. To be more precise, consider the joint likelihood of $\boldsymbol{x}_t$, $\boldsymbol{y}_t$ and m_t[2]:

$$\begin{aligned} & p(X, Y, \boldsymbol{m}|\Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega, \beta, \sigma_\nu, \boldsymbol{b}) \\ =& p(Y|\Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega) \\ & \cdot p(\boldsymbol{m}|Y, \Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega, \beta, \sigma_\nu, \boldsymbol{b}) \\ & \cdot p(X|\boldsymbol{m}, Y, \Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega) \end{aligned} \tag{3.13}$$

where $X = [\boldsymbol{x}_1, ..., \boldsymbol{x}_T]$, $Y = [\boldsymbol{y}_1, ..., \boldsymbol{y}_T]$ and $\boldsymbol{m} = [m_1, ..., m_T]$ stack the observational series, the factors, and the instrument horizontally. Note that, while the marginal likelihood of Y and the conditional likelihood of $X|\boldsymbol{m}, Y$ depend only on reduced form parameters, the conditional likelihood of $\boldsymbol{m}|Y$ depends, in addition, on structural parameters, β, σ_ν and $\boldsymbol{b}$.

[2]The factors are identified only up to an invertible rotation, i.e. the representations $\boldsymbol{x}_t = \Lambda \boldsymbol{y}_t + \boldsymbol{\xi}_t$ and $\boldsymbol{x}_t = \Lambda P P^{-1} \boldsymbol{y}_t + \boldsymbol{\xi}_t$ yield the same likelihood. Therefore, in order to achieve identification, one has to impose further restrictions. I follow Bernanke et al. (2005) and set the upper $R \times K$ block of Λ^z equal to a zero matrix. Furthermore, $\Sigma^f = Cov(\boldsymbol{f}_t)$ is diagonal and $\frac{\Lambda^{f'}\Lambda^f}{N} = I_R$. This normalisation, although not necessary, is sufficient to pin down the factor rotation, as pointed out by Kilian and Lütkepohl (2017).

The conditional likelihood of $\boldsymbol{m}|Y$ can be written as (see Appendix 3.A.2 for a derivation):

$$\boldsymbol{m}|Y \sim N(\boldsymbol{\mu}_{m|Y}, V_{m|Y}) \tag{3.14}$$

$$\boldsymbol{\mu}_{m|Y} = \beta\boldsymbol{\epsilon}_1 \tag{3.15}$$

$$V_{m|Y} = \sigma_\nu^2 I_T, \tag{3.16}$$

where $\boldsymbol{\epsilon}_1 = [\epsilon_{1,1}, ..., \epsilon_{1,T}]$ stacks the structural shocks of interest in a vector.

As seen in equation (3.15), the conditional likelihood of $\boldsymbol{m}$ is higher the higher the correlation between $\boldsymbol{m}$ and $\beta\boldsymbol{\epsilon_1}$. In the posterior sampler, draws are weighted by the conditional likelihood of $\boldsymbol{m}$ (see Appendix 3.A.3). Therefore, the econometrician will give more weight to posterior draws, which result in structural errors that look like a scaled version of the proxy.

As is apparent from equation (3.15), since $\boldsymbol{\epsilon}_1$ is obtained from reduced form errors, identification depends heavily on the model specification. Therefore, one should pay close attention to which variables are included in the model since an omitted variable bias translates into biases in the identified structural shocks. Augmenting the model with latent factors helps alleviate this problem without taking a stand on which of a potentially large set of observational series need to be included.

Compared to recursively identified VARs, the BP-FAVAR has the advantage that when using a Proxy VAR, the researcher is not forced to employ potentially non-credible short run exclusion restrictions. For example, the recursive workhorse model on the empirical identification of uncertainty shocks by Bloom (2009) relies on the assumption that financial markets do not price an exogenous increase in uncertainty within a month. This might be too strong an assumption given the fast-moving nature of financial markets. In the context of Proxy VAR models such short run exclusion restrictions are substituted by external information contained in the proxy variable, m_t.

3.2.3 Inference

The Bayesian approach treats all model parameters and latent factors as random variables whose posterior needs to be sampled from. In order to outline the sampling procedure, first define the parameter space as

$$\theta = (\Pi, \Sigma, \Lambda^f, \Lambda^z, \Omega, \beta, \sigma_\nu, \boldsymbol{b}) \tag{3.17}$$

The joint posterior of parameters and latent factors is:

$$p(\theta, F|X, Z, \boldsymbol{m}), \tag{3.18}$$

where $F = [\boldsymbol{f}_1, ..., \boldsymbol{f}_T]$, $Z = [\boldsymbol{z}_1, ..., \boldsymbol{z}_T]$. The challenge consists in approximating the marginal posterior distributions of the latent factors,

$$p(F|X, Z, \boldsymbol{m}) = \int_\theta p(\theta, F|X, Z, \boldsymbol{m}) d\theta \tag{3.19}$$

and the model parameters,

$$p(\theta|X, Z, \boldsymbol{m}) = \int_F p(\theta, F|X, Z, \boldsymbol{m}) dF \tag{3.20}$$

It is shown by Geman and Geman (1984) that these integrals can be approximated using a multi-move Gibbs sampler, which alternately draws from two distributions: First, draw the latent factors given all model parameters and the data, i.e.

$$p(F|\theta, X, Z, \boldsymbol{m}). \tag{3.21}$$

This draw is generated using filtering techniques. Second, draw the model parameters conditioning on this draw of factors and the data, i.e.[3]

$$p(\theta|Z, X, \boldsymbol{m}). \tag{3.22}$$

This draw is generated using a Metropolis-within-Gibbs algorithm as in Caldara and Herbst (2019).

Compared to the common approach of first extracting factors via Principal Components and then feeding them into a VAR (see Stock and Watson, 2016 for a review), this Bayesian approach has the advantage that it allows for Bayesian shrinkage of the parameter space. This might seem unnecessary given that the factors already reduce the dimensionality of the estimation problem. However, if, as in the present case, the number of observable factors, K, or the lag length, P, is large, dimensionality issues still arise and can be alleviated using Bayesian shrinkage.[4] Furthermore, as shown by

[3] I follow Caldara and Herbst (2019) and generate draws from $p(F|\theta, X)$ and $p(\theta|Z, X)$ first and account for the additional conditioning on $\boldsymbol{m}$ using an independence Metropolis-Hastings step. See Appendix 3.A.3 for details.

[4] In the application, the number of parameters to be estimated in the transition equation is

$$\begin{aligned}&(R+K)(1+(R+K)P)+(R+K)^2\\ =&(4+8)(1+(4+8)5)+(4+8)(4+8+1)/2\\ =&810.\end{aligned}$$

Yamamoto (2019), bootstrap inference in frequentist factor models is far from trivial. In particular, it remains an open issue how to account for the estimation uncertainty in the factors. A Bayesian approach, on the other hand, offers a unified way of summarising the uncertainty of the model, as pointed out also by Huber and Fischer (2018). The joint posterior summarises estimation uncertainty in both the parameters and the latent factors.

Conditional posterior densities of latent factors F given θ: The procedure to generate posterior draws of latent factors, F, differs from generating draws of parameters, θ, in that one has to generate the whole dynamic evolution of factors for each $t = 1, ..., T$. For this to be feasible I exploit the Markov property of the system described in equation (3.3) as follows:

$$p(Y|X,\theta) = p(\boldsymbol{y}_t|X,\theta)\prod_{t=1}^{T-1} p(\boldsymbol{y}_t|\boldsymbol{y}_{t+1},X,\theta). \tag{3.23}$$

First note that (3.23) describes the posterior of Y, which contains both latent and observable factors. The reason for including the observable factors is the dynamic interdependence between latent and observable factors, which needs to be accounted for. Given that the observable factors are non-random, their distribution has a zero variance.[5] Second, note that this is a product of $R + K$-dimensional conditional distributions. Given the assumption of Gaussianity of $\boldsymbol{\xi}_t$ and $\boldsymbol{u}_t$, this representation can be combined with the observation equation (3.1) and is amenable to the Carter-Kohn algorithm described in Carter and Kohn (1994) and Frühwirth-Schnatter (1994) (see Appendix 3.A.5 for details). This approach, while straightforward to implement, increases the computational burden slightly compared to Principal Components Analysis. However, it allows incorporating the estimation uncertainty in the latent factors in a consistent way.

Conditional posterior densities of the parameters $\boldsymbol{\theta}$ given latent factors $\boldsymbol{F}$: In order to draw the model parameters θ given the data and a draw of the factors, Y, I form three blocks of parameters: Block 1 refers to parameters of the observation equation (3.1), block 2 refers to the parameters of the transition equation (3.3), and block 3 refers to the parameters of the proxy equation (3.7). Conditional on a draw of the factors and the data the first two blocks can be sampled independently from each other while the last block is sampled conditional on the second block.

Block 1: Observation Equation The idiosyncratic errors $\boldsymbol{\xi}_t$ are assumed to be mutually uncorrelated and normally distributed. Group the factor loading matrices as

The sample length is $T = 438$.

[5]Here, I refer to the variance across draws. The variance across time is, of course, non-zero.

$\Lambda = [\Lambda^f \quad \Lambda^z]$. Then, conditional on a draw of the factors, we can specify conjugate normal-inverse Gamma priors and draw the posterior for Λ and Ω equation-by-equation using well-known results on Bayesian linear regression models (see e.g. Koop, 2003). For each equation i, specify the priors as:

$$\boldsymbol{\omega}_{ii} \sim IG(sc^*, sh^*) \tag{3.24}$$

$$\boldsymbol{\lambda}_i|\boldsymbol{\omega}_{ii} \sim N(\boldsymbol{\mu}^*_{\lambda,i}, \boldsymbol{\omega}_{ii}M_i^{*-1}), \tag{3.25}$$

where $\boldsymbol{\lambda}_i$ is the i-th row of Λ and $\boldsymbol{\omega}_{ii}$ is the i-th diagonal element of Ω. These priors translate into posterior distributions of the following form:

$$\boldsymbol{\omega}_{ii}|X, Y \sim IG(\bar{sc}_i, \bar{sh}_i) \tag{3.26}$$

$$\lambda_i|\boldsymbol{\omega}_{ii}, X, Y \sim N(\bar{\mu}_{\lambda,i}, \boldsymbol{\omega}_{ii}\bar{M}_i^{-1}) \tag{3.27}$$

with

$$\bar{sh}_i = sh^* + T \tag{3.28}$$

$$\bar{sc}_i = sc^* + \hat{\boldsymbol{\xi}}_{\boldsymbol{i}}\hat{\boldsymbol{\xi}}'_{\boldsymbol{i}} + (\hat{\boldsymbol{\lambda}}_i^{OLS} - \mu^*_{\boldsymbol{\lambda},i})'(\bar{M}_i^{-1} + (YY')^{-1})^{-1}(\hat{\boldsymbol{\lambda}}_i^{OLS} - \mu^*_{\boldsymbol{\lambda},i}) \tag{3.29}$$

$$\hat{\boldsymbol{\lambda}}_i^{OLS} = \boldsymbol{x}_i Y'(YY')^{-1} \tag{3.30}$$

$$\hat{\boldsymbol{\xi}}_{\boldsymbol{i}} = \boldsymbol{x}_i - \hat{\boldsymbol{\lambda}}_i^{OLS} Y \tag{3.31}$$

$$\bar{M}_i = M_i^* + YY' \tag{3.32}$$

$$\bar{\mu}_{\boldsymbol{\lambda},i} = \bar{M}_i(M_i^{*-1}\mu^*_{\boldsymbol{\lambda},i} + YY'\hat{\boldsymbol{\lambda}}_i^{OLS}) \tag{3.33}$$

I follow Bernanke et al. (2005) in specifying $sc^* = 3$, $sh^* = 10^{-3}$, $M_i^* = I_{R+K}$ and $\mu^*_{\boldsymbol{\lambda},i} = \mathbf{0}_{(R+K)\times 1}$.

Block 2: Transition Equation Given a draw of factors, $\boldsymbol{y}_t$ follows a standard $VAR(P)$ model. Therefore, we can employ a version of the Minnesota/ Litterman prior (Litterman, 1986) and specify independent normal-inverse Wishart priors:

$$vec(\Pi) \sim N(\boldsymbol{\mu}^*_\Pi, V^*_\Pi) \tag{3.34}$$

$$\Sigma \sim IW(S^*, \tau^*), \tag{3.35}$$

where $vec(\cdot)$ is the vectorisation operator that stacks the column of a matrix one underneath the other into a vector. These priors translate into the following conditional

posterior for $vec(\Pi)$:

$$vec(\Pi)|\Sigma, Y \sim N(\bar{\boldsymbol{\mu}}_\Pi, \bar{V}_\Pi) \tag{3.36}$$

$$\bar{V}_\Pi = (V_\Pi^{*-1} + (WW' \otimes \Sigma^{-1}))^{-1} \tag{3.37}$$

$$\bar{\boldsymbol{\mu}}_\Pi = \bar{V}_\Pi(V_\Pi^{*-1}\boldsymbol{\mu}_\Pi^* + (W \otimes \Sigma^{-1})vec(Y)) \tag{3.38}$$

where $W = [\boldsymbol{w}_1, ..., \boldsymbol{w}_T]$, $\boldsymbol{w}_t = [\mathbf{1} \quad \boldsymbol{y}_{t-1} \quad ... \quad \boldsymbol{y}_{t-p}]'$ stacks a vector of 1s and P lags of $\boldsymbol{y}_t$. The conditional posterior for Σ is given as:

$$\Sigma|\Pi, Y \sim IW(\bar{S}, \bar{\tau}) \tag{3.39}$$

$$\bar{S} = S^* + UU' \tag{3.40}$$

$$\bar{\tau} = \tau^* + T, \tag{3.41}$$

where $U = [\boldsymbol{u}_1, ..., \boldsymbol{u}_T]$ stacks the reduced form errors given the current draw of factors. I set $\boldsymbol{\mu}_\Pi^*$ to a zero vector given that all series are transformed to be stationary. V_Π^* is a diagonal matrix containing the prior variances of the parameters contained in Π. These are set in accordance with standard Minnesota values and given as:

$$v_{\Pi,i,j}^* = \begin{cases} (\lambda/l)^2 & \text{if} \quad i = j \\ (\lambda\sigma_i/l\sigma_j)^2 & \text{if} \quad i \neq j, \end{cases} \tag{3.42}$$

where σ_i is obtained from univariate AR(1) regressions and $\lambda = 0.2$. S^* is set to I_{R+K}, while τ^* is set to $R + K$.

Block 3: Proxy Equation The parameters of the proxy equation are sampled conditional on the parameters of the transition equation. For β and σ_ν the priors are

$$\beta \sim N(\mu_\beta^*, \sigma_\beta^*) \tag{3.43}$$

$$\sigma_\nu \sim IG(sc_\nu^*, sh_\nu^*). \tag{3.44}$$

$\boldsymbol{b}$ is computed as $\boldsymbol{b} = chol(\Sigma)Q_{\cdot,1}$ where $Q_{\cdot,1}$ is the first column of a draw from the uniform Haar distribution (see Rubio-Ramirez et al., 2010 for a discussion).

These priors translate into the following posteriors for β and σ_ν:

$$\beta|Y, \boldsymbol{m}, \Pi, \Sigma, \sigma_\nu, \boldsymbol{b} \sim N(\bar{\mu}_\beta, \bar{\sigma}_\beta) \tag{3.45}$$

$$\bar{\mu}_\beta = \boldsymbol{m}\boldsymbol{\epsilon_1}(\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_1}')^{-1} \tag{3.46}$$

$$\bar{\sigma}_\beta = \sigma_\nu(\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_1}')^{-1} \tag{3.47}$$

$$\sigma_\nu | Y, \boldsymbol{m}, \Pi, \Sigma, \beta, \boldsymbol{b} \sim IG(\bar{sc}_\nu, \bar{sh}_\nu) \tag{3.48}$$

$$\bar{sc}_\nu = sc^*_\nu + (\boldsymbol{m} - \beta\boldsymbol{\epsilon_1})(\boldsymbol{m} - \beta\boldsymbol{\epsilon_1})' \tag{3.49}$$

$$\bar{sh}_\nu = sh^*_\nu + T \tag{3.50}$$

For $\boldsymbol{b}$ the conditional posterior has an unknown form. Therefore, $\boldsymbol{b}$ is sampled using a Metropolis-Hastings step. In particular, at iteration j, a draw $\boldsymbol{b}^{cand}$, will be accepted with probability (see Appendix 3.A.3 for details)

$$\alpha = min(\frac{p(\boldsymbol{m}|Y, \Pi, \Sigma, \boldsymbol{b}^{cand})}{p(\boldsymbol{m}|Y, \Pi, \Sigma, \boldsymbol{b}^{j-1})}, 1) \tag{3.51}$$

I follow Caldara and Herbst (2019) in specifying the priors as $\mu^*_\beta = 0$, $\sigma^*_\beta = 1$, $sh^*_\nu = 2$, $sc^*_\nu = 0.2$ in order to allow the data to dominate the posterior. In particular, this prior specification implies a zero mean prior correlation between instrument and shock, ρ.

Posterior Sampler The sampler can be summarized as follows (see the Appendix 3.A.3 for a detailed step-by-step procedure):

1. Set starting values
2. Draw $\boldsymbol{y}_t$ via the Carter-Kohn algorithm
3. Draw Λ and Ω
4. Draw Σ and Π
5. Draw $\boldsymbol{b}$ using a Metropolis-Hastings step
6. Draw β and σ_ν

These steps are repeated a sufficient number of times for the algorithm to converge (see Appendix 3.A.5 for a discussion of the convergence properties of the sampler).

3.3 Data, Estimation and Results

This section first describes the data and the proxy used for identification. It then shows how the number of factors is determined with the goal of alleviating informational insufficiency issues. Lastly, it discusses how instrument relevance is assessed in the Bayesian context and presents impulse responses from the baseline model as well as from three benchmark models.

3.3.1 Data and Transformations

The baseline data contained in $\boldsymbol{z}_t$ are monthly US data from Piffer and Podstawski (2017), who updated the data in Bloom (2009). The vector of observational series, $\boldsymbol{x}_t$, contains 126 of the monthly FRED dataset by McCracken and Ng (2016) that are not already included in $\boldsymbol{z_t}$ (see Appendix 3.A.7 for a detailed description).[6] The length of the estimation sample is constrained by the instrument[7] and lasts from 1979M1 through 2015M7.[8] I follow Piffer and Podstawski (2017) in setting the lag length to $P = 5$ as a baseline (see Appendix 3.A for a robustness exercise setting $P = 9$).

One concern is the measurement of the monetary policy stance. Traditionally, the effective federal funds rate is considered to be the policy tool of the central bank. However, given that it was constrained by the zero lower bound in the period 2009M1 to 2015M11, this variable cannot serve as an indicator of the policy stance. This is why, for this period, I replace the effective federal funds rate by the shadow rate as computed by Wu and Xia (2016). It is based on a term structure model and often considered a better reflection of the policy stance during the zero lower bound period than the federal funds rate. Figure 3.16 in Appendix 3.A.7 shows the shadow rate.

All informational series contained in $\boldsymbol{x}_t$ are transformed to have zero mean. In addition, they are transformed to induce stationarity as proposed by McCracken and Ng (2016). Missing values are replaced with zeros, which is the unconditional mean of the standardized series. These missing values occur mostly at the beginning of the dataset and amount to less than one percent of total observations. Therefore, the joint dynamics are unlikely to be overly affected by this imputation.

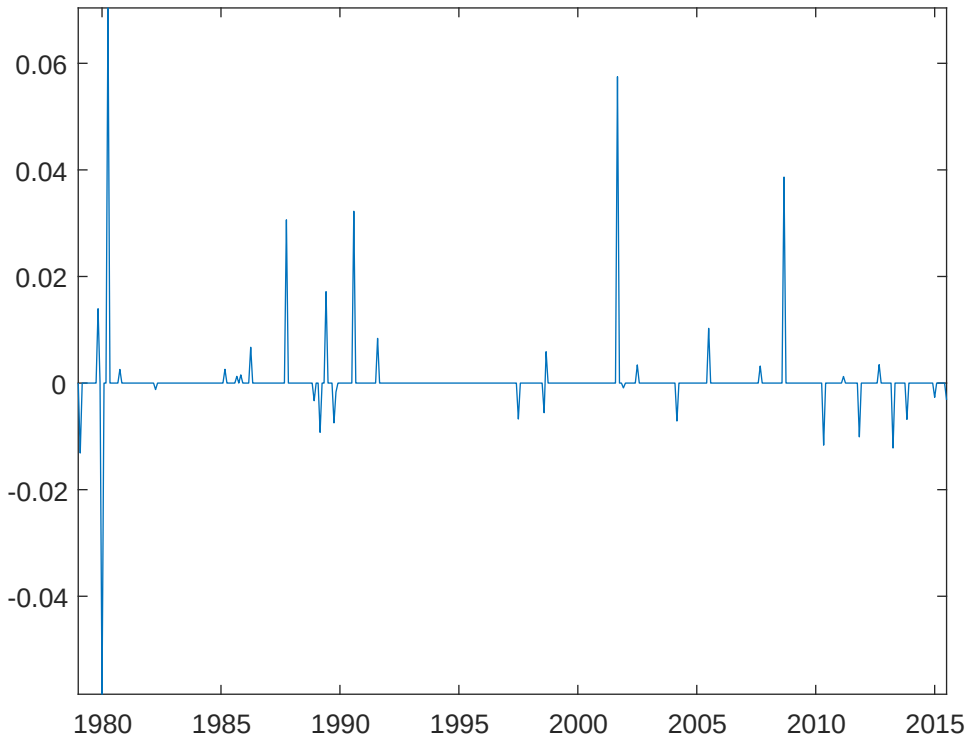
Piffer and Podstawski (2017) argue that a proxy for the uncertainty shock could be based on the price of gold. The intuition behind this idea is that gold is considered a safe haven asset that investors choose in times of heightened uncertainty. This generates movements in the price of gold. The challenge consists in finding price variations that are not correlated with structural shocks other than the uncertainty shock. In order to achieve this, the authors collect a series of 38 events, which are considered to be associated with movements in uncertainty (e.g. the fall of the Berlin Wall or the 9/11 terrorist attacks). They then compute the variation of the gold price in narrow windows around these events and argue that these variations are driven exclusively by movements in uncertainty. In addition, they show that this proxy has a low correlation

[6] Data set available at https://research.stlouisfed.org/econ/mccracken/fred-databases/

[7] Instrument available at https://sites.google.com/site/michelepiffereconomics/home/research-1

[8] An alternative to shortening the sample is put forward by Braun et al. (2017). They suggest generating synthetic observations for the instrument within their posterior sampler. This procedure implicitly requires parameter stability for time periods when the instrument is unavailable and it assumes that missing values occur at random.

Figure 3.1: Uncertainty Proxy



Note: Gold price variation around selected events. The sample period is 1979M1 to 2015M7.

with other structural shocks as computed by Stock and Watson (2012), which is further evidence for exogeneity to their system.

Figure 3.1 shows the proxy. It peaks during well-known events such as the 9/11 terrorist attacks in 2001 or the bankruptcy of Lehman brothers in 2008.

3.3.2 Determining the Number of Factors

Choosing the number of latent factors, R, has important consequences for the amount of additional information the BP-FAVAR is based on, compared to the small-scale VAR model employed in Bloom (2009) and Piffer and Podstawski (2017). I follow Mandalinci and Mumtaz (2019) and base the choice of the number of factors on the criterion proposed by Bai and Ng (2002). Table 3.1 shows the criterion (see Appendix 3.A.6 for a description). It suggests to set $R = 4$. An alternative criterion for choosing the number of factors is a scree plot (see Appendix 3.A.6 for a discussion and additional results).

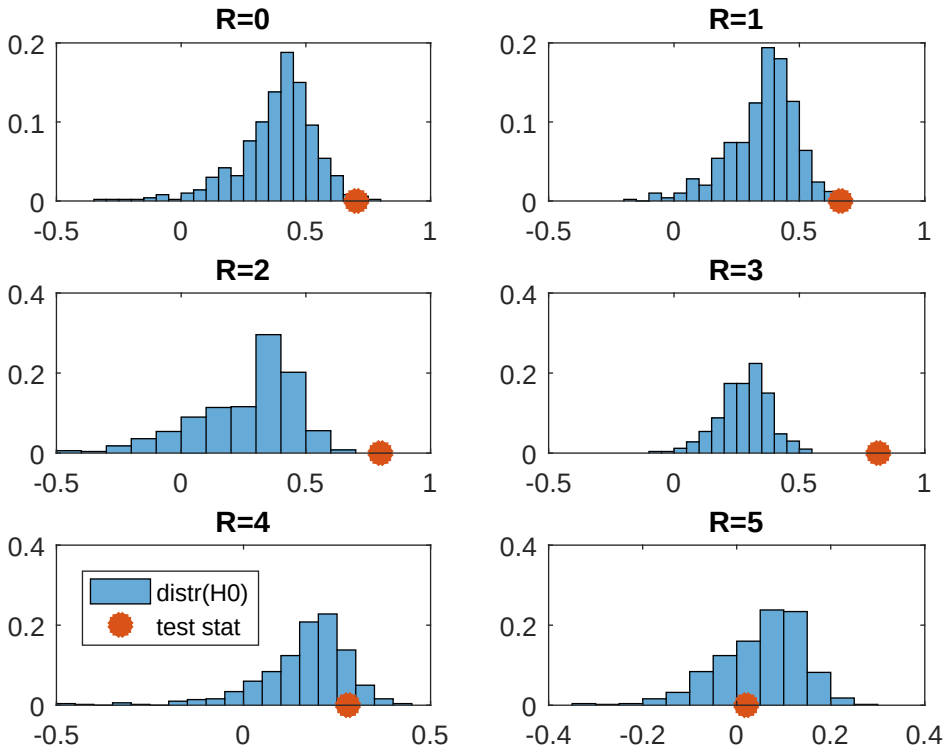
In order to present some evidence that including $R = 4$ aligns the econometrician's and the economic agent's information sets, I employ the sequential testing procedure suggested by Forni and Gambetti (2014). Figure 3.2 shows the test statistic under the

Table 3.1: Bai and Ng (2002) criterion

R	Criterion
1	-0.1736
2	-0.3033
3	-0.4163
4	**-0.4767**
5	-0.4504
6	-0.4168
7	-0.3777
8	-0.3329
9	-0.2868
10	-0.2398

Note: See Appendix 3.A.6 for a description.

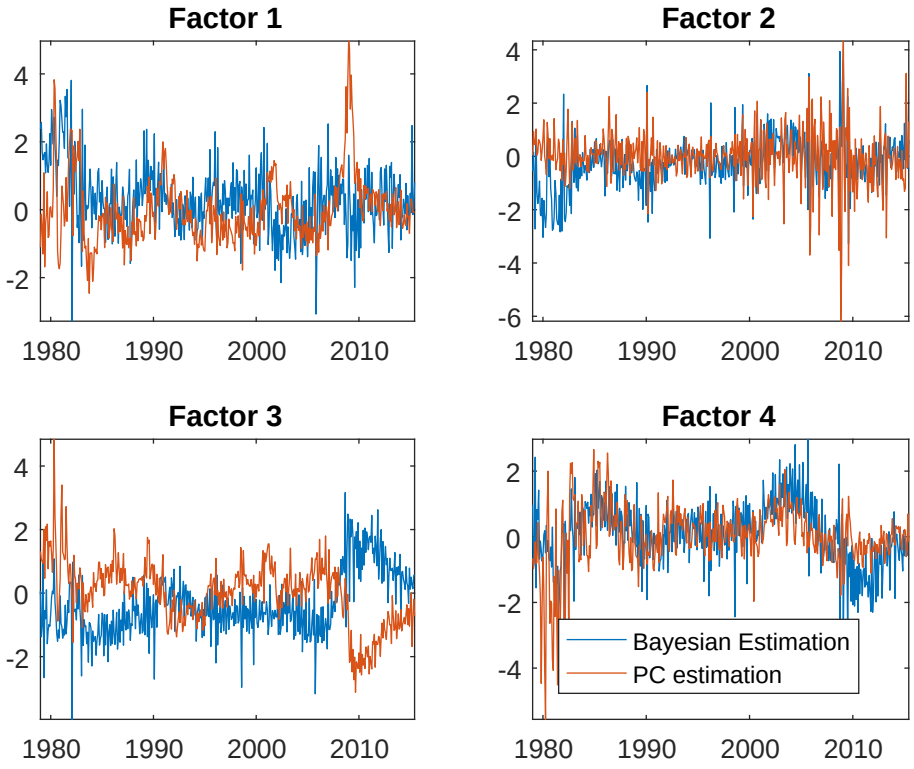
Figure 3.2: Test for Informational Sufficiency



Note: The histogram shows the bootstrap test statistic under the Null of no Granger-causality. It is based on a multivariate one-step-ahead out-of-sample Granger-causality test with lag length 4. The bootstrap test statistic is based on 1000 replications. The sample is split as $T = T_1 + T_2$, $T_1 = T_2 = 0.5T$. A larger test statistic indicates Granger-causality. Rejection of the Null of no Granger-causality indicates informational deficiency.

Null of informational sufficiency together with the actual test statistic (see Appendix 3.A.4 for a detailed description). The test for $R = 0$ can be considered evidence that the workhorse model of Bloom (2009) is, indeed, informationally deficient and needs to be extended in order to alleviate potential omitted variables biases. The Null Hypothesis of informational sufficiency cannot be rejected once the model is augmented by at least four factors. This suggest that setting $R = 4$ alleviates informational deficiency issues. Figure 3.3 shows the last accepted draw of estimated factors from the Bayesian

Figure 3.3: Factors



Note: Last accepted draw from posterior sampler and Principal Components estimation of factors

estimation, together with the Principal Components estimation used as a starting value.

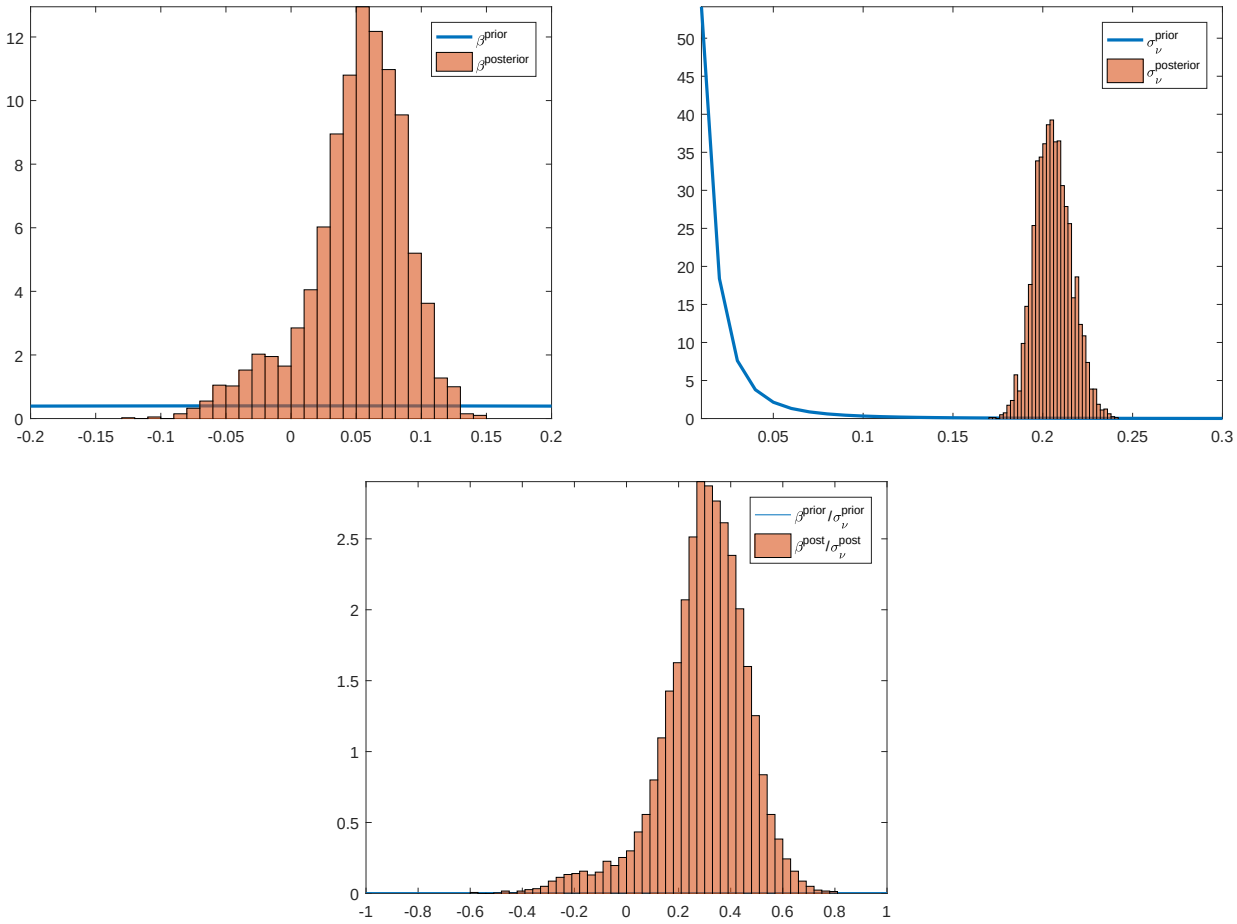
3.3.3 Relevance of the Instrument

The instrument is appropriate to identify the uncertainty shock to the extent that it contains enough information about the shock, i.e. it is relevant. The relevance of the instrument in the Bayesian context is assessed by analysing the posterior updating of the proxy equation. A high posterior correlation between instrument and shock suggests relevance.

Figure 3.4 shows the updating of the relevant quantities β, σ_ν and their ratio β/σ_ν. This ration is the signal-to-noise ratio and measures how much information the instrument contains about the shock of interest.

The top left panel shows that while using a prior for β that is flat over the relevant parameter space, the posterior is centred around 0.1 suggesting that the data support a correlation between structural error $\epsilon_{1,t}$ and instrument m_t. The top right panel shows the updating of σ_ν. The prior is chosen to have mean 0.02 and infinite variance, as in Caldara and Herbst (2019). The posterior suggests a standard deviation of this noise measurement centred around 0.2. The bottom panel shows the implications for the signal-to-noise ratio. While the prior is centred around zero and flat over the whole parameter space, the posterior is centred around 0.15 and has little probability mass near zero. This strongly suggests that the instrument contains relevant information about the structural shock.

Figure 3.4: Instrument Relevance



Note: Updating of β (top left), σ_ν (top right) and β/σ_ν (bottom panel). The blue line shows the prior distribution while the histogram shows draws from the posterior distribution.

3.3.4 Updating of the impact effects

Given that both the recursive identification scheme and the proxy identification scheme impose structure on the impact effect of shocks, differences between these two approaches will be most apparent in the identification of $\boldsymbol{b}$. It contains the impact effects of an uncertainty shock on the latent and observable factors. The prior distribution is not available in closed form but is implicit in the prior distributions of Σ, $Q_{\cdot,1}$, β and σ_ν. Prior draws are generated imposing the prior mean for β, i.e. setting $\beta = 0$, so that all rotation vectors, $Q_{\cdot,1}$, are accepted with equal probability. A draw from the prior of $\boldsymbol{b}$ conditional on the factors is computed as follows:

- Draw Σ^{prior} from its prior inverse Wishart distribution
- Draw $Q_{\cdot,1}^{prior}$ as the first column of a draw from the uniform Haar distribution
- Compute $\boldsymbol{b}^{prior} = chol(\Sigma)Q_{\cdot,1}$.

As pointed out by Baumeister and Hamilton (2015), a uniform prior on $Q_{\cdot,1}$ does not necessarily imply a uniform distribution over the structural parameters of interest, which in this case are the elements of $\boldsymbol{b}$. Figure 3.5 shows that, indeed, the implicit prior on $\boldsymbol{b}$ has some curvature on the impact effect of the S&P 500 index and the VXO. However, it has good coverage of the relevant parameter space. More importantly, it is flat in the relevant parameter regions for the variables of prime interest, namely CPI, hours, employment, and industrial production as well as the latent factors. Therefore, the implicit prior on $\boldsymbol{b}$ should not overly affect the posterior of $\boldsymbol{b}$.

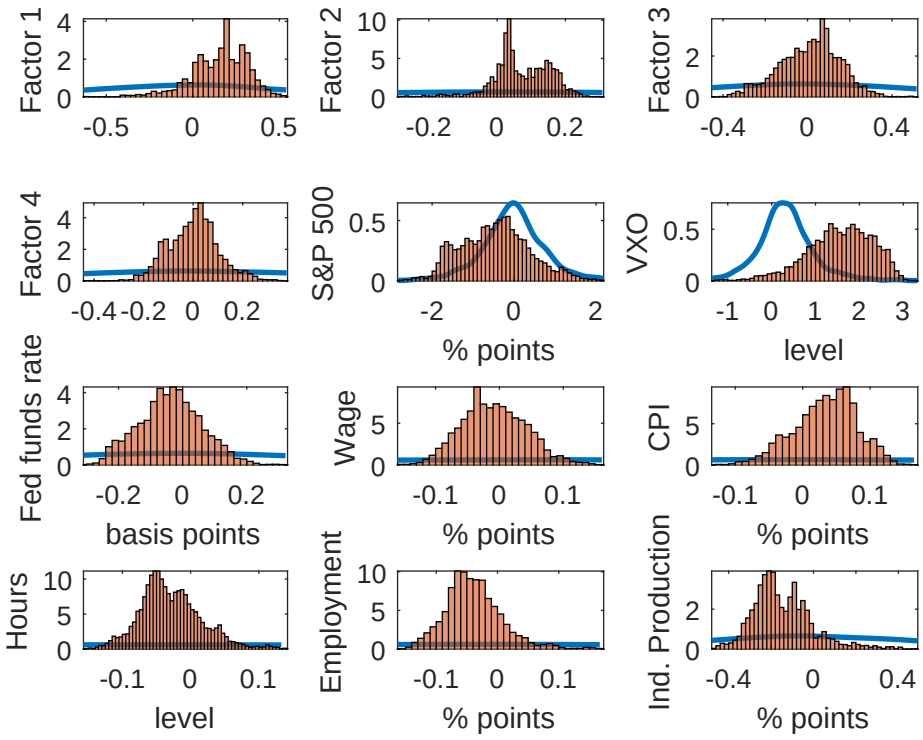
3.3.5 Impulse Responses

The BP-FAVAR extends the workhorse model by Bloom (2009) in two ways: First, instead of imposing exclusion restrictions on $\boldsymbol{b}$, the BP-FAVAR achieves identification via a proxy. Second, the BP-FAVAR addresses informational deficiency of the Bloom (2009) model, which could potentially bias the responses of variables to the uncertainty shock. The workhorse model contains the following variables: $\Delta log(S\&P500)$, VXO, federal funds rate, $\Delta log(wages)$, $\Delta log(CPI)$, hours, $\Delta log(employment)$, and $\Delta log(IP)$. In order to isolate the effects of the identification scheme and the information set, I include three benchmark models to compare to the baseline BP-FAVAR:

BP-VAR. This model can be considered a Bayesian adaptation of the model by Piffer and Podstawski (2017). I use their model specification, i.e.

$$\boldsymbol{y}_t = \boldsymbol{z}_t, \tag{3.52}$$

Figure 3.5: Updating of b



Note: Updating of the first column of B (not normalised). The blue line shows the prior distribution of $\boldsymbol{b}$ computed as the distribution implicit in the priors on Σ and $Q_{\cdot,1}$ and plotted using a Kernel smoother. The bars show the posterior.

and identification is achieved via the gold price proxy. Differences between the BP-FAVAR and the BP-VAR are driven primarily by informational issues.

Recursively identified VAR. This model is akin to Bloom (2009). For consistency the variable selection is the same as in the previous model, but identification is achieved by imposing a lower-triangular structure on B, i.e.

$$B = chol(\Sigma). \tag{3.53}$$

The uncertainty shock is ordered second, i.e. after the S&P500 index. This assumes that the stock market does not react within a month to an exogenous increase in uncertainty and that financial shocks are the only shocks, apart from the uncertainty shock itself, which influence the uncertainty measure within the month. Differences between the BP-FAVAR and the recursive VAR can be driven by both informational and identification issues.

Recursively identified FAVAR. This scheme adds R latent factors to the small-scale VAR while keeping B lower-triangular. The uncertainty shock is ordered in position $R+2$, i.e. after the latent factors and the S&P500 index. This assumes that neither the latent factors nor the stock market react within a month to an exogenous increase in uncertainty. Differences between the BP-FAVAR and the recursive FAVAR in the impact effects of shocks will be driven primarily by the identification scheme.

Figure 3.6 shows the impulse responses obtained from the baseline and the three benchmark models: The response variables are those employed by Bloom (2009), Piffer and Podstawski (2017) and other studies. The shock is normalized to generate an increase of 2.5 in the VXO on impact, which is comparable in magnitude to these studies. Estimation is based on 50000 Gibbs draws, discarding the first 40000 draws as a burn-in sample, as in Belviso and Milani (2006) and Amir-Ahmadi and Uhlig (2015).

All four models replicate the main findings of Bloom (2009): A rapid drop and subsequent rebound of employment, production and hours worked. It is also in line with Basu and Bundick (2017) who argue that the co-movement among these variables is a key empirical feature that theoretical models should be able to reproduce.

For the stock market, the two recursively identified models (column 3 and 4) exclude an on-impact effect of the uncertainty shock on the stock market index. This results in a biased reaction of the S&P 500 index: The models identified via a proxy (columns 1 and 2) show that the stock market reacts on impact and quickly rebounds. This is in line with the fast moving nature of financial markets, which price increases in uncertainty within the period.

For the real variables, the recursive scheme suggests a moderate negative reaction of hours, employment and industrial production of less than -0.05%. In the two models identified via proxies this reaction is estimated to be up to -0.2% for industrial production and roughly -0.1% for hours and employment. Given the similarity in the price reaction across models, the recursive models provide a skewed view of the nominal and real interactions following an uncertainty shock.

The inclusion of factors does not qualitatively alter the results. The BP-FAVAR produces results broadly in line with the BP-VAR. The biases resulting from a recursive identification in the small-scale VAR cannot be alleviated by the inclusion of factors. As pointed out by Sims (2012), informational insufficiency is not an either/or concept but can have quantitatively very different effects on impulse responses, depending on the application. In the present case, statistical tests detect informational insufficiency, but alleviating this issue does not have a severe quantitative impact.

3.4 Conclusion

This paper aims at recovering the interrelations between real and nominal variables following an identified uncertainty shock by employing a Bayesian Proxy FAVAR.

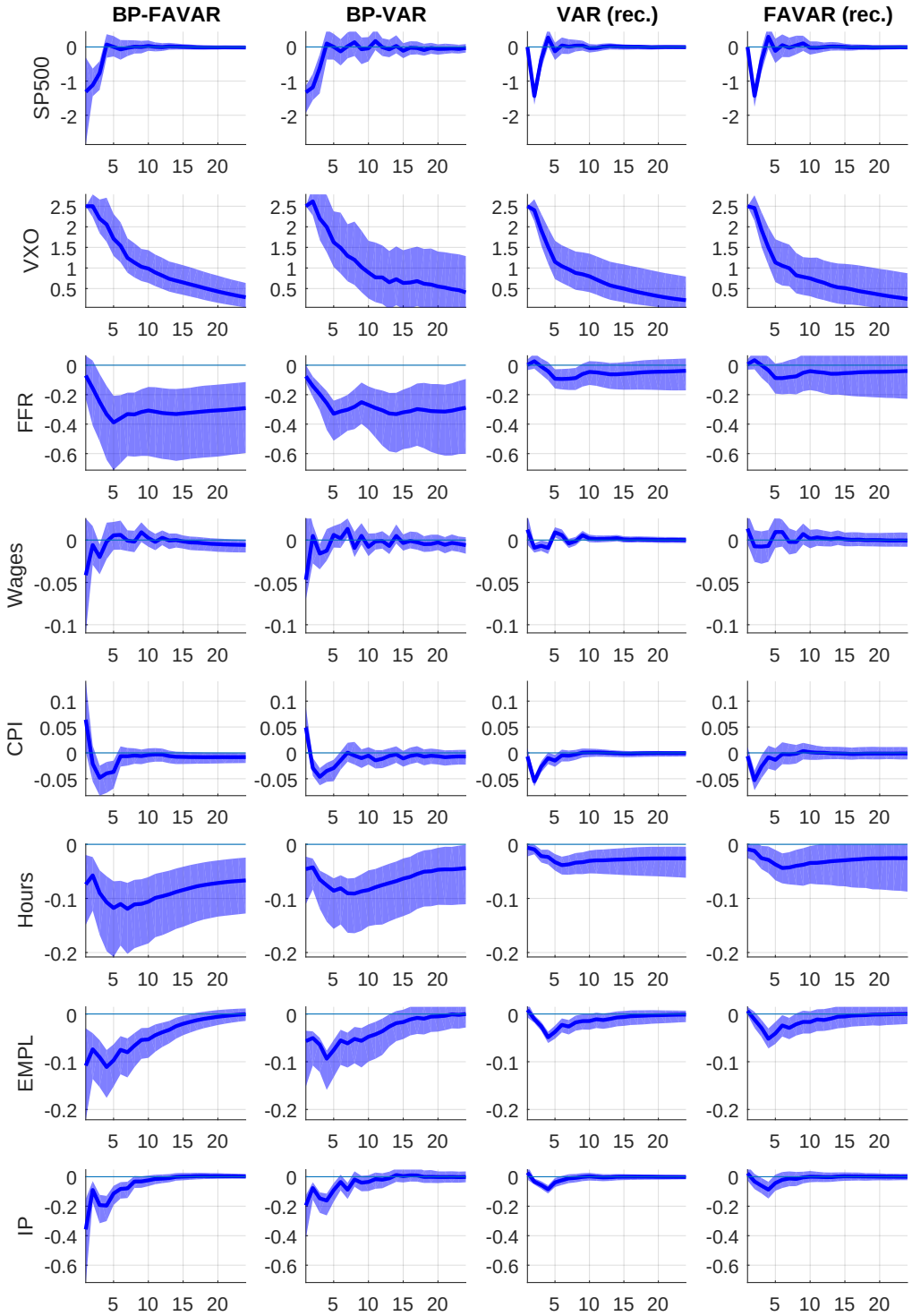
The first contribution is the empirical finding that uncertainty shocks are deflationary approximately two months after the impact and indistinguishable from zero after six months. This finding relates to the work by Born and Pfeifer (2014) and Fernández-Villaverde et al. (2015) who show in a theoretical modelling framework that the inflationary effects of uncertainty shocks are driven by two opposing channels: an aggregate demand and a price-bias channel. The aggregate effect of these two channels is unknown ex ante. My results lend empirical support to the dominance of the aggregate demand over the price-bias channel.

The second contribution is methodological. I combine a recent strand of the Bayesian VAR literature that uses external instruments for identification (Caldara and Herbst, 2019) with the Bayesian factor model literature (Bernanke et al., 2005 Belviso and Milani, 2006, Amir-Ahmadi and Uhlig, 2015). I show how a state-space model can be set up to jointly exploit the advantages of both approaches. The resulting Bayesian Proxy factor-augmented VAR model avoids two shortcomings of commonly employed small-scale recursively identified VAR models, namely a non-credible identification scheme and informational insufficiency. I detect informational insufficiency of the small-scale workhorse model, but find that it has limited quantitative impact on the estimated impulse responses to an uncertainty shock. This relates to the work by Sims (2012) who also finds that informational deficiency is not an either/or but a quantitative issue.

Future empirical research concerned with the identification of uncertainty shocks in a structural VAR context, especially if it is conducted with few variables, should pay close attention to the information set. Informational insufficiency is detected even in the relatively rich workhorse model by Bloom (2009). While it has only limited quantitative impacts in this case, this might be very different when reducing the information set.

From a methodological point of view, the BP-FAVAR model offers potential for a number of extensions: First, identification of two or more shocks is generally possible in this set-up. As pointed out by Kilian and Lütkepohl (2017), avoiding non-credible short-run exclusion restrictions is particularly important in factor models. Therefore, the BP-FAVAR could be extended to identify multiple shocks via external instruments. Second, a combination of proxies with sign restrictions in this context is a natural point of departure given the similarity in model set-up between the BP-VAR and VARs identified via sign restrictions. The combination of these two approaches is likely to lead to sharper inference.

Figure 3.6: Impulse Responses



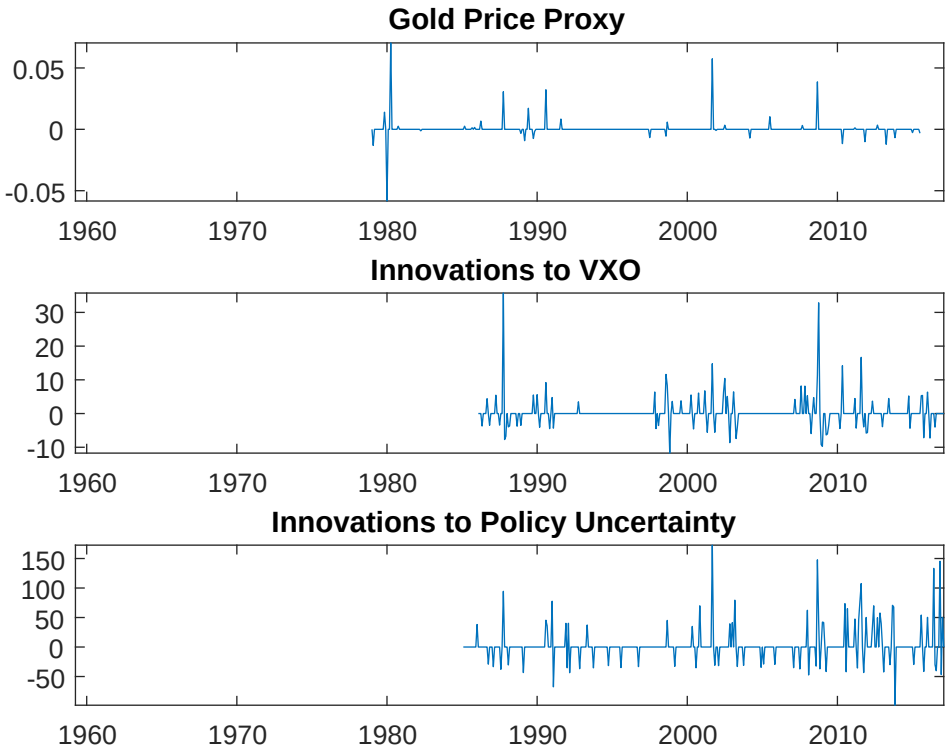
Note: The sample is 1979M1 - 2015M7. The first column shows IRFs from the baseline BP-FAVAR. The second column shows IRFs from a small-scale Proxy VAR. The third column shows IRFs from a small-scale recursively identified VAR. The last column shows IRFs from a recursively identified FAVAR. The BP-FAVAR and the recursive FAVAR are based on $R = 4$ latent factors. The bands are computed point-wise 68 % posterior credible bands based on 50000 draws of the posterior sampler discarding the first 40000 as burn-in.

3.A Further Results

3.A.1 Robustness Checks

Stock and Watson (2012) were the first to propose two proxies for uncertainty shocks. They employ the innovations in the VXO and in the policy uncertainty index by Baker et al. (2016). These innovations are computed as residuals from an AR(2) process. While innovations to the VXO are likely to be correlated with the uncertainty shocks of interest in this study, this is less likely for the policy uncertainty instrument. The reason is that the concept of uncertainty it is designed to approximate differs from the one I have in mind in this study. Thus, in analysing impulse responses, I concentrate on innovations to the VXO.

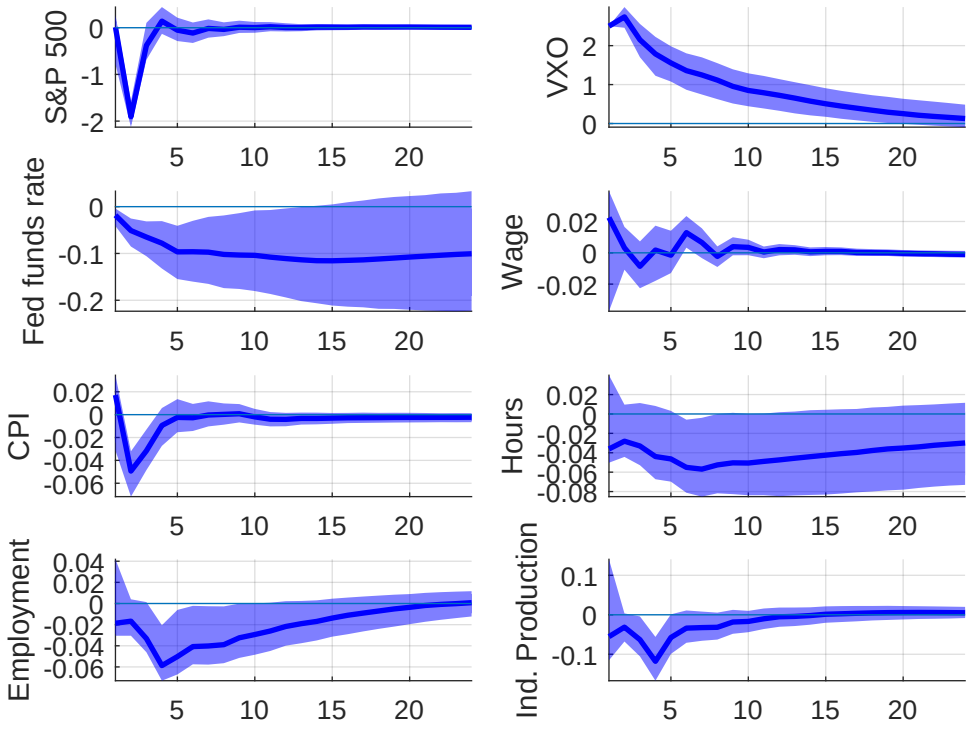
Figure 3.7: Uncertainty Proxies



Note: The top panel shows the baseline gold price proxy for uncertainty by Piffer and Podstawski (2017). The middle panel shows the Stock and Watson (2012) uncertainty proxy computed as residuals of an AR(2) process of the VXO. The bottom panel shows the Stock and Watson (2012) uncertainty proxy computed as residuals of an AR(2) process of the Baker et al. (2016) policy uncertainty proxy. All proxies are winsorized at the 10 % level.

I reproduce the instrument as follows: In a first step I estimate an AR(2) process including a constant for the VXO index as in Stock and Watson (2012). In a second

Figure 3.8: Stock and Watson (2012) innovations to the VXO



Note: The model is specified as in the baseline with the gold price proxy replaced by the Stock and Watson (2012) VXO innovations. These are computed as the residuals from an AR(2) process of the VXO.

step, I truncate the proxy at the 10th and 90th percentile and set all observations within these two quantiles to zero. This is to ensure that the proxy only captures quantitatively relevant spikes in uncertainty.

Figure 3.7 plots the resulting proxies while Table 3.2 shows the correlations among the proxies. The instruments have some degree of correlation, but they clearly measure distinct variations in uncertainty.

Table 3.2: Correlation among proxies

Gold proxy	1.00		
VXO innovations	0.38	1.00	
Policy Uncertainty innovations	0.50	0.37	1.00

Figure 3.8 shows the impulse response functions when employing the innovations to the VXO as proxies. While the identification of the impact effects is not as sharp as in the baseline, the dynamics are broadly similar.

In order to assess robustness with respect to the lag length, I set $P = 9$. Figure 3.9 shows that the main results remain unchanged.

Jurado et al. (2015) argue that indices like the VXO used in the baseline do not accurately capture the type of macroeconomic uncertainty that the researcher is interested in. They propose alternative indeces based on large panels of time series. Jurado et al. (2015) provide measures extracted from macro series and Ludvigson et al. (2018) provide measures extracted from financial and real series.[9] Figure 3.10 shows that they are highly correlated over the whole sample period.

Figure 3.11 replaces the VXO by the macro uncertainty index, Figure 3.12 replaces the VXO by the financial uncertainty index and Figure 3.13 replaces it by the real uncertainty index. The results remain qualitatively unchanged.

3.A.2 Conditional likelihood of m_t

This section re-parametrises Caldara and Herbst (2019) to allow for identification of impact effects. For the posterior sampler, we will need to be able to evaluate the conditional likelihood of m_t given $\boldsymbol{y}_t$.

Restate the model for convenience:

$$\boldsymbol{x}_t = \Lambda \boldsymbol{y}_t + \boldsymbol{\xi}_t \tag{3.54}$$

$$\begin{bmatrix} \boldsymbol{y}_t \\ m_t \end{bmatrix} = \begin{bmatrix} \Pi & 0 \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_t \\ m_{t-1} \end{bmatrix} + \begin{bmatrix} B & \boldsymbol{\beta} \\ \boldsymbol{\beta}' & \sigma_\nu \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \nu_t \end{bmatrix} \tag{3.55}$$

$$Var(\begin{bmatrix} \boldsymbol{y}_t - \Pi \boldsymbol{w}_t \\ m_t \end{bmatrix}) = \begin{bmatrix} \boldsymbol{\Sigma} & B\boldsymbol{\beta}' \\ \boldsymbol{\beta} B' & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_\nu^2 \end{bmatrix}, \tag{3.56}$$

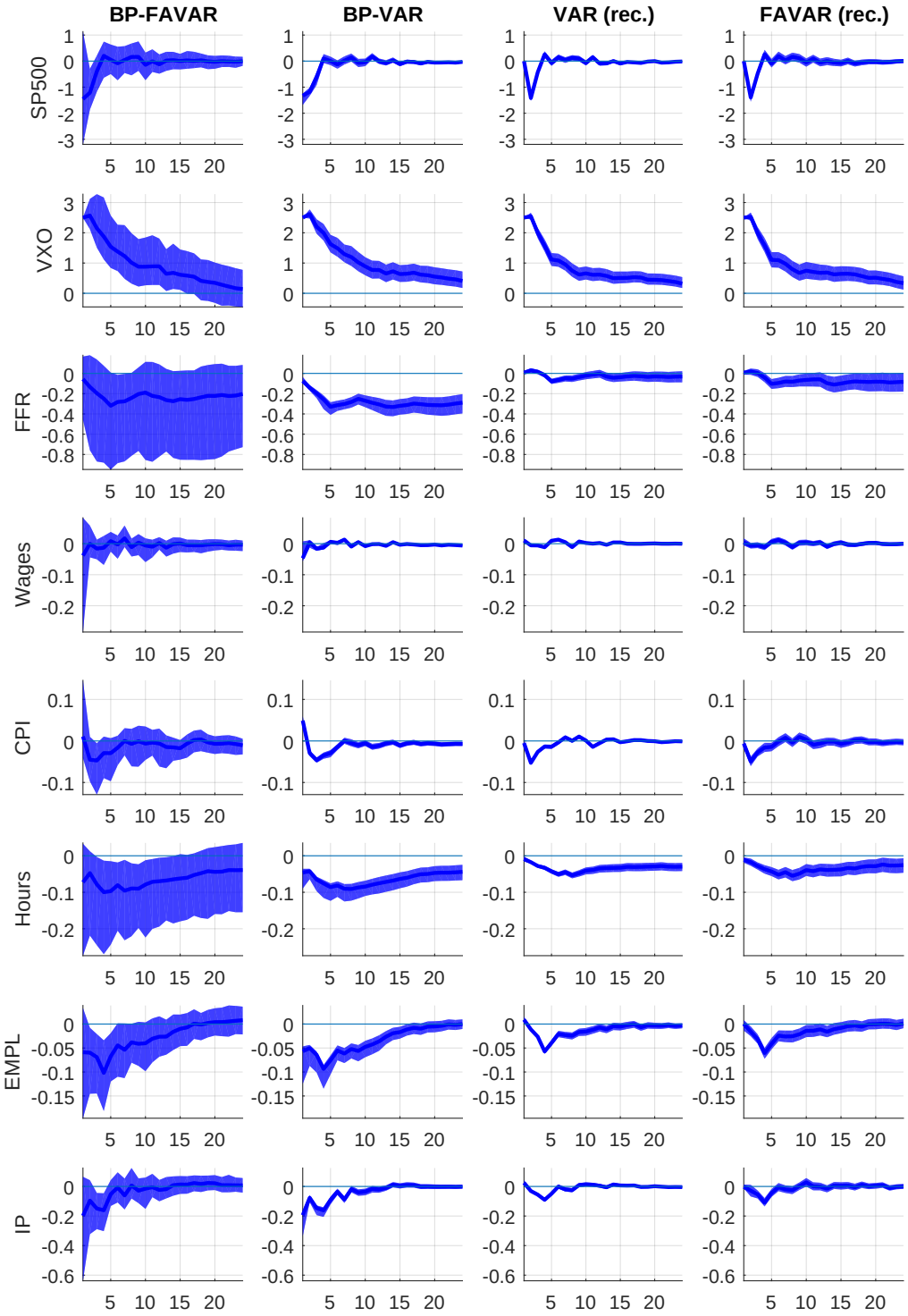
where $\boldsymbol{\beta} = [\beta \quad \mathbf{0}]'$

The likelihood is invariant to observationally equivalent rotations of B. Therefore we can replace $B = B^c Q$, where B^c is, for example, the lower-triangular Cholesky decomposition of Σ.

$$Var(\begin{bmatrix} \boldsymbol{y}_t - \Pi \boldsymbol{w}_t \\ m_t \end{bmatrix}) = \begin{bmatrix} \Sigma & B^c Q \boldsymbol{\beta}' \\ \boldsymbol{\beta} Q' B^{c'} & \boldsymbol{\beta}'\boldsymbol{\beta} + \sigma_\nu^2 \end{bmatrix} \tag{3.57}$$

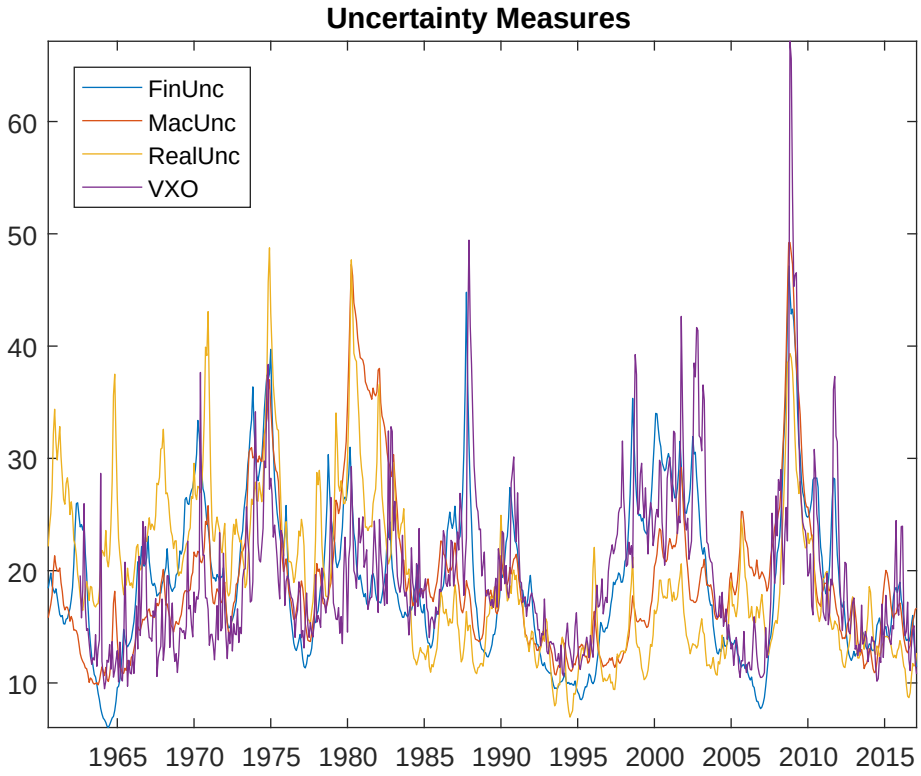
[9] Details can be found here https://www.sydneyludvigson.com/data-and-appendixes/

Figure 3.9: Impulse Responses (P=9)



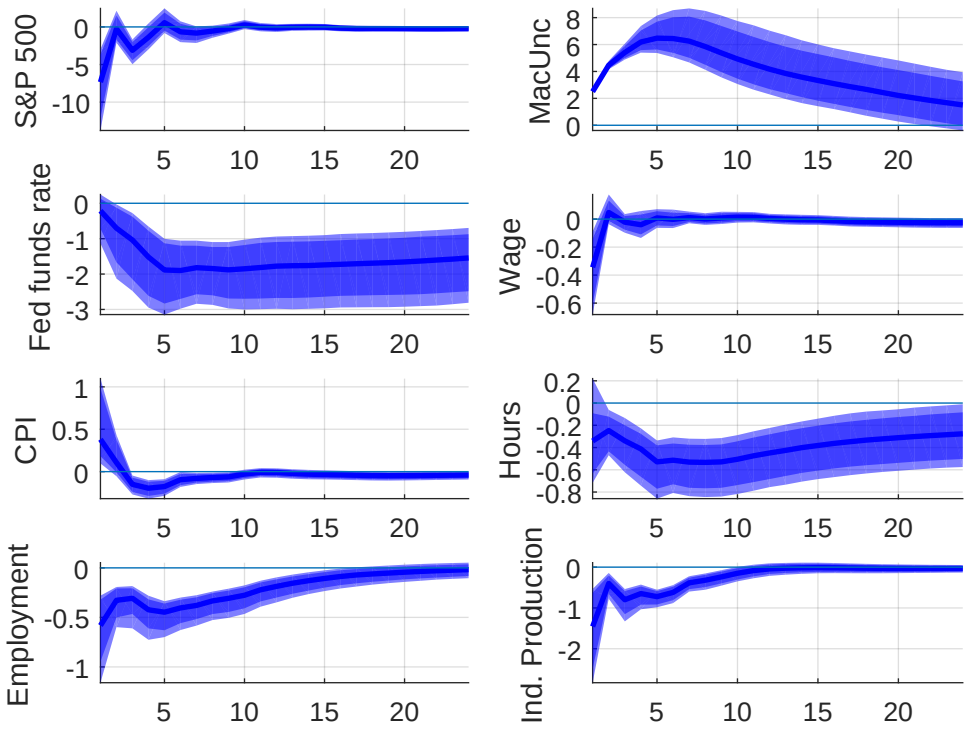
Note: The sample is 1979M1 - 2015M7. The first column shows IRFs from the baseline BP-FAVAR. The second column shows IRFs from a small-scale Proxy VAR. The third column shows IRFs from a small-scale recursively identified VAR. The last column shows IRFs from a recursively identifed FAVAR. The bands are computed point-wise 68 % posterior credible bands based on 10000 draws of the posterior sampler.

Figure 3.10: Uncertainty Measures



Note: The uncertainty measures are taken from Ludvigson et al. (2018) and Ludvigson et al. (2018) and rescaled to match the mean of the VXO for comparability.

Figure 3.11: Macroeconomic Uncertainty



Note: The model is specified as in the baseline with VXO replaced by the macroeconomic uncertainty indicator by Ludvigson et al. (2018).

Then, using the rules for the conditional mean of multivariate normal distributions, we obtain the conditional likelihood

$$m_t|\boldsymbol{y}_t, \Pi, \Sigma, \boldsymbol{b}, \beta, \sigma_\nu \sim N(\mu_{m|Y}, V_{m|Y}), \tag{3.58}$$

$$\mu_{m|Y} = \boldsymbol{\beta} Q' B^{c'} \Sigma^{-1} \boldsymbol{u}_t \tag{3.59}$$

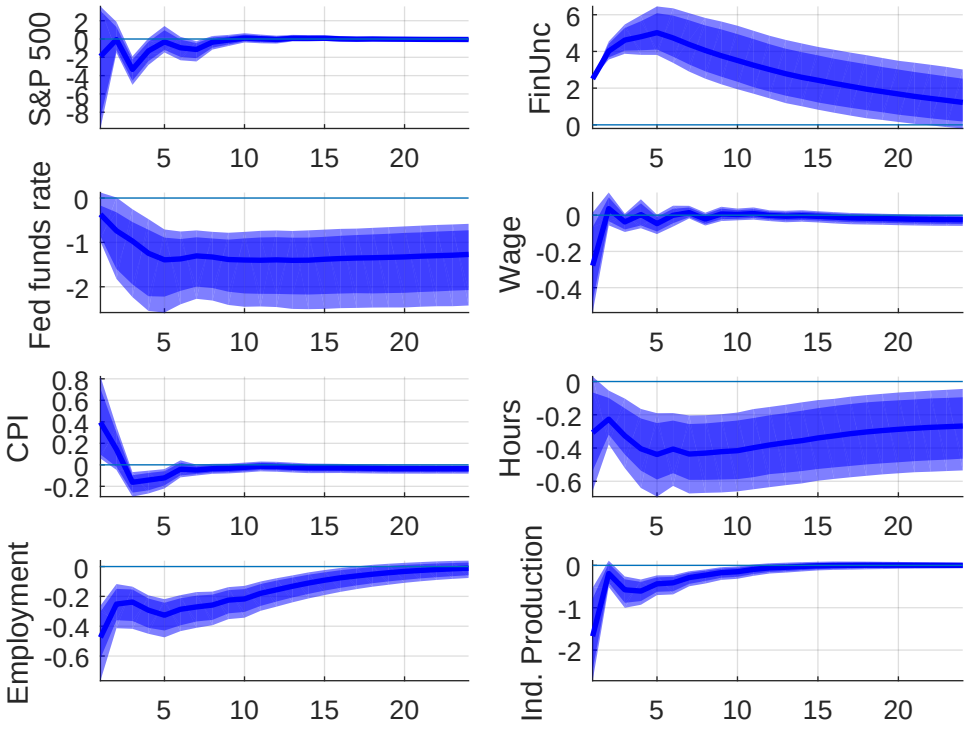
$$= \beta \boldsymbol{\epsilon}_{1,t} \tag{3.60}$$

$$V_{m|Y} = \boldsymbol{b}\boldsymbol{b}' + \sigma_\nu^2 - \boldsymbol{b} Q' B^{c'} \Sigma^{-1} B^c Q \boldsymbol{b}' \tag{3.61}$$

$$= \sigma_\nu^2 \tag{3.62}$$

Note that, once we condition on $\boldsymbol{y}_t$, the likelihood of m_t does not depend on $\boldsymbol{x}_t$. In addition, note that the conditional likelihood of m_t does not depend on the full matrix B, but only on its first column, $\boldsymbol{b}$ because the model is partially identified.

Figure 3.12: Financial Uncertainty



Note: The model is specified as in the baseline with VXO replaced by the financial uncertainty indicator by Ludvigson et al. (2018).

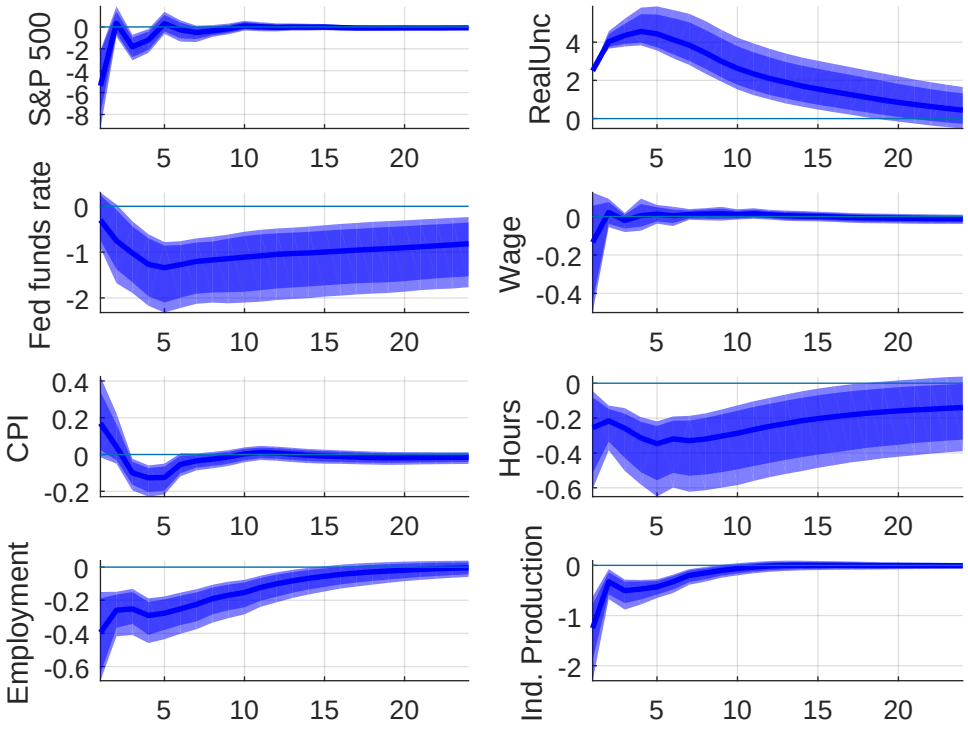
3.A.3 Metropolis-within-Gibbs sampler for the BP-FAVAR

This section outlines the posterior sampling procedure for the Bayesian Proxy Factor Augmented VAR. It combines posterior samplers for Bayesian FAVARs, e.g. Koop et al. (2010) and Amir-Ahmadi and Uhlig (2015) with the algorithm for Bayesian Proxy VARs proposed by Caldara and Herbst (2019). It provides posterior draws of the parameters $[\Pi, \Sigma, \Omega, \Lambda, \beta, \sigma_\nu]$ as well as of the latent factors $\boldsymbol{f}_t$.

Start by rewriting the posterior distribution as

$$\begin{aligned} p(\theta, Y|X, \boldsymbol{m}) &\propto p(X, \boldsymbol{m}|\theta, Y)p(\theta, Y) \\ &= p(\boldsymbol{m}|\theta, Y, X)p(X|\theta, Y)p(\theta, Y) \\ &= p(\boldsymbol{m}|\theta, Y, X)p(X, \theta, Y) \\ &= p(\boldsymbol{m}|\theta, Y, X)p(\theta, Y|X)p(X) \\ &\propto p(\boldsymbol{m}|\theta, Y, X)p(\theta, Y|X) \\ &\propto p(\boldsymbol{m}|\theta, Y)p(\theta, Y|X) \end{aligned}$$

Figure 3.13: Real Uncertainty



Note: The model is specified as in the baseline with VXO replaced by the real uncertainty indicator by Ludvigson et al. (2018).

Note that the second term in the last line is the posterior distribution of all model parameters and factors given the data. Standard results (see Amir-Ahmadi and Uhlig, 2015) exist to generate draws from this posterior. The first term is the conditional likelihood of $\boldsymbol{m}|Y$. Note that the last transformation is justified by the fact that once conditioned on Y, X does not contain further information about $\boldsymbol{m}$. The posterior sampler weights draws from $p(\boldsymbol{\theta}, Y|X)$ with the conditional likelihood of $\boldsymbol{m}|Y$ using an independence Metropolis-Hastings step as in Caldara and Herbst (2019). It is in this sense that $\boldsymbol{m}$ informs the estimation of reduced form parameters.

1. Set starting values

   In order to obtain starting values of the reduced form parameters $[\Pi, \Sigma, \Omega, \Lambda]$, I estimate the model once using the two-step-procedure proposed by Boivin et al. (2009), which takes the restrictions implied by the observation equation into account when extracting factors. I use Principal Component Analysis to obtain R factors $\boldsymbol{f}_t^{PC}$ from $\boldsymbol{x}_t$ and the factor loadings Λ^{PC}. Then run the regression

$$\boldsymbol{x}_t = const + \Lambda^f \boldsymbol{f}_t^{PC} + \Lambda^z \boldsymbol{z}_t + \boldsymbol{v}_t$$

and construct $\tilde{X} = \hat{\Lambda}^f \boldsymbol{f}_t^{PC}$, the fitted values orthogonalized with respect to the observable factors. Then extract R factors from $\tilde{X}$ and repeat the procedure 20 times as in Boivin et al. (2009). Save $[\Lambda^0, \boldsymbol{f}_t^0, \Omega^0]$

Lastly, estimate a reduced-form VAR in $\boldsymbol{y}_t^0 = \begin{bmatrix} \boldsymbol{f}_t^0 \\ \boldsymbol{z}_t \end{bmatrix}$ to obtain $[\Sigma^0, \Pi^0]$.

For the remaining parameters, I start the algorithm from $\beta^0 = 0$, $\sigma_\nu^0 = 0.5std(m_t)$ (Caldara and Herbst, 2019 call this the "high relevance prior", which imposes that half the variance in the proxy can be attributed to measurement error)

At each stage j proceed with the following steps:

2. Draw $\boldsymbol{f}_t^j$ using the Carter-Kohn backward recursion of the Kalman filter. Set $\boldsymbol{y}_t^j = \begin{bmatrix} \boldsymbol{f}_t^j \\ \boldsymbol{z}_t \end{bmatrix}$ (see Appendix 3.A.5 for details.)

3. Draw Λ^j from its conditional normal posterior given in equation (3.27).

$$\lambda_i | \boldsymbol{\omega}_{ii}^{j-1}, X, Y \sim N(\bar{\boldsymbol{\mu}}_{\lambda,i}, \boldsymbol{\omega}_{ii}^{j-1} \bar{M}_i^{-1})$$

Impose the normalisation on Λ.

4. Compute $\boldsymbol{\xi}_t^j = \boldsymbol{x}_t - \Lambda^j \boldsymbol{y}_t^j$ and draw the diagonal elements of Ω from their posterior inverse Gamma distributions

$$\boldsymbol{\omega}_{ii} | X, Y \sim IG(\bar{sc}_i/2, \bar{sh}_i/2)$$

5. Sample Σ^{cand} from an inverse Wishart

$$\Sigma^{cand} \sim IW(\bar{S}, \bar{\tau})$$

6. Sample Π^{cand} from a multivariate normal using the Minnesota values as priors for the posterior covariance and employ shrinkage towards a white noise process given the stationarity of the system.

$$vec(\Pi^{cand}) \sim N(\bar{\boldsymbol{\mu}}_\Pi, \bar{V}_\Pi)$$

7. With probability α set $\Pi^j = \Pi^{cand}$ and $\Sigma^j = \Sigma^{cand}$, otherwise set $\Pi^j = \Pi^{j-1}$ and $\Sigma^j = \Sigma^{j-1}$, where

$$\alpha = min(\frac{p(\boldsymbol{m}, Y | \Pi^{cand}, \Sigma^{cand}, Q^{j-1})}{p(\boldsymbol{m}, Y | \Pi^{j-1}, \Sigma^{j-1}, Q^{j-1})}, 1)$$

8. Draw $Q_{\cdot,1}^{cand}$ as the first column of an orthogonal matrix form a uniform Haar distribution using the algorithm by Rubio-Ramirez et al. (2010). Set $Q_{\cdot,1}^{j} = Q_{\cdot,1}^{cand}$ with probability α and $Q_{\cdot,1}^{j} = Q_{\cdot,1}^{j-1}$ else.

$$\alpha = min(\frac{p(\boldsymbol{m}|Y, \Pi^j, \Sigma^j, Q_{\cdot,1}^{cand})}{p(\boldsymbol{m}|Y, \Pi^j, \Sigma^j, Q_{\cdot,1}^{j-1})}, 1)$$

9. Compute structural errors $\epsilon_t^j = (chol(\Sigma^j)Q_{\cdot,1}^j)^{-1}U^j$. Draw β^j from its posterior normal distribution

$$\beta^j \sim N(\bar{\mu}_\beta, \bar{\sigma}_\beta)$$

10. Draw σ_ν^j from its posterior inverse Gamma distribution

$$\sigma_\nu^j \sim IG(\bar{sh}, \bar{sc})$$

3.A.4 Test for Informational Sufficiency

In this section I describe the sequential testing procedure to assess whether the BP-FAVAR model aligns the econometrician's and the agent's information set. This would not be the case in the presence of omitted variables. This procedure is particularly suited when the list of potentially omitted variables is large and the researcher therefore would like to avoid taking a stand on which additional variables to include. In the context of uncertainty shocks, variables which are typically omitted from small-scale VAR models but could potentially be relevant, include consumer sentiment (Bachmann and Sims, 2012), total factor productivity (Bachmann and Bayer, 2013), and measures of anticipated risk (Christiano et al., 2014). Other forward-looking variables could also potentially cause omitted variable biases.

Instead of including one variable after the other in the VAR the sequential testing procedure augments the small-scale VAR model by latent factors extracted from a large set of informational series until the model is informationally sufficient. The model is informationally sufficient if none of the observable variables is Granger-caused by factors. The basis is a multivariate out-of-sample Granger-causality test (see next section for details). The intuition behind this test is the following: If the economy is accurately represented by a factor model, as is assumed here, then the factors contain all relevant information that agents base their decision making on. If these factors do not help predict a vector of variables in the VAR, then the variables in the VAR contain the same information as the factors. Thus, they are sufficient to align the

econometrician's and the agent's information set. If, on the other hand, the factors help predict the VAR variables, then they should be subsequently added to the VAR as additional variables until informational sufficiency cannot be rejected any longer. The sequential testing procedure is as follows:

In a first step, I test

$$H_0 : \boldsymbol{f}_t^{PC} \text{ do not Granger-cause } \boldsymbol{z}_t \tag{3.63}$$

$$H_1 : \boldsymbol{f}_t^{PC} \text{ Granger-cause } \boldsymbol{z}_t, \tag{3.64}$$

where $\boldsymbol{f}_t^{PC}$ is a vector containing the first six Principal Components extracted from $\boldsymbol{x}_t$. I then test, for $j = 1, ..., 5$

$$H_0 : \boldsymbol{f}_{t,-(1:j)}^{PC} \text{ do not Granger-cause } \{\boldsymbol{z}_t, \boldsymbol{f}_{t,1:j}\} \tag{3.65}$$

$$H_1 : \boldsymbol{f}_{t,-(1:j)}^{PC} \text{ Granger-cause } \{\boldsymbol{z}_t, \boldsymbol{f}_{t,1:j}\}. \tag{3.66}$$

where $\boldsymbol{f}_{t,1:j}^{PC}$ are the first j Principal Components of $\boldsymbol{x}_t$ and $\boldsymbol{f}_{t,-(1:j)}^{PC}$ is a vector containing all but the first j Principal Components.

Figure 3.2 in the paper shows the distribution of the test statistic across 1000 samples obtained via a standard residual bootstrap procedure (see next section for details) together with the test statistic computed using the actual sample data. If the actual test statistic lies outside the bootstrap distribution test statistic, then this indicates that the Null of no Granger causality can be rejected. Table 3.3 shows the corresponding p-values. It suggests that for $R = 0$, informational sufficiency can be rejected. The Null Hypothesis of informational sufficiency cannot be rejected once the model is augmented by at least four factors. Therefore, the multiple testing procedure suggests using $R = 4$.

Table 3.3: Test for Informational Sufficiency

R	**p-value**
0	0.0000
1	0.0020
2	0.0000
3	0.0000
4	0.1240
5	0.6680

Note: p-values are obtained as the fraction of bootstrap test statistics under the Null of no Granger causality exceeding the actual test statistic.

Multivariate out-of-sample Granger-causality test

We would like to assess whether a vector $\boldsymbol{f}_t$ Granger-causes a vector $\boldsymbol{z}_t$. The series $\boldsymbol{f}_t$ is said to Granger-cause the series $\boldsymbol{z}_t$ if the past of $\boldsymbol{f}_t$ has additional power for forecasting $\boldsymbol{z}_t$ after controlling for the past of $\boldsymbol{z}_t$. Gelper and Croux (2007) base their test statistic on the comparison of two nested VAR models:

$$\boldsymbol{z}_t = \Phi(L)\boldsymbol{z}_{t-1} + \boldsymbol{v}_t^r \tag{3.67}$$

$$\boldsymbol{z}_t = \Phi(L)\boldsymbol{z}_{t-1} + \Psi(L)\boldsymbol{f}_{t-1} + \boldsymbol{v}_t^f \tag{3.68}$$

The restricted model (3.67) has only past values of $\boldsymbol{z}_t$ as regressors, while the unrestricted model (3.68) has both the past of $\boldsymbol{z}_t$ and $\boldsymbol{f}_t$ as regressors. $\boldsymbol{v}_t^r$ are the residuals of the restricted, while $\boldsymbol{v}_t^f$ are the residuals of the full model. The test statistic is based on the out-of-sample forecast performance of these two models. Compared to in-sample tests, this approach is less susceptible to overfitting.

The unrestricted model can be written out as

$$\boldsymbol{z}_t = \phi_0 + \phi_1\boldsymbol{z}_{t-1} + ... + \phi_p\boldsymbol{z}_{t-p} + \psi_1\boldsymbol{f}_{t-1} + ... + \psi_p\boldsymbol{f}_{t-p} + \boldsymbol{v}_t^f \tag{3.69}$$

where ϕ_j, $j = 0, ..., p$ are of dimension $K \times K$ and ψ_j are of dimension $K \times R$. Then the Null hypothesis of no Granger causality can be written as

$$H_0 : \psi_1 = \psi_2 = \cdots = \psi_p = 0 \tag{3.70}$$

The out-of-sample test proceeds as follows: First, split the sample in half as $T = T_1 + T_2$, where $T_1 = T_2 = 0.5T$ (assuming T is even) and construct one-step-ahead forecasts $\hat{\boldsymbol{z}}_{T_1+1}^r$ and $\hat{\boldsymbol{z}}_{T_1+1}^f$ based on the restricted and the full model, respectively. Then, expand the estimation sample by one and construct $\hat{\boldsymbol{z}}_{T_1+2}^r$ and $\hat{\boldsymbol{z}}_{T_1+2}^f$. The last forecasts, $\hat{\boldsymbol{z}}_T^r$ and $\hat{\boldsymbol{z}}_T^f$ are based on an estimation sample of size $T-1$. As a second step, construct the series of one-step-ahead forecast errors $\hat{\boldsymbol{v}}_t^r = \hat{\boldsymbol{z}}_t^r - \boldsymbol{z}_t$ and $\hat{\boldsymbol{v}}^f = \hat{\boldsymbol{z}}_t^f - \boldsymbol{z}_t$ for the two competing models and save them in the vectors $\hat{\boldsymbol{v}}^r$ and $\hat{\boldsymbol{v}}^f$ of size T_2. As a third step, construct the test statistic comparing the forecasting performance of the two models as

$$MSFE = log(\frac{|\hat{\boldsymbol{v}}^{r\prime}\hat{\boldsymbol{v}}^r|}{|\hat{\boldsymbol{v}}^{f\prime}\hat{\boldsymbol{v}}^f|}), \tag{3.71}$$

where MSFE is the mean squared forecast error and $|\cdot|$ stands for the determinant of a matrix. If the full model provides better forecasts, MSFE takes a larger value, indicating Granger-causality.

The asymptotic distribution of the test statistic is unknown. Critical values will therefore be based on a residual bootstrap. It proceeds in the following steps:

1. Estimate the model under the Null, i.e. model (3.67) and compute the test statistic, denoted here by s_0, as described above
2. Generate $Nb = 1000$ new time series $\boldsymbol{z}_1^*, ..., \boldsymbol{z}_t^*$ according to model (3.67) using the parameter estimates and resampling the residuals with replacement. For each bootstrap sample, compute the test statistic resulting in $s_1^*, ..., s_{Nb}^*$
3. The percentage of bootstrap test statistics, $s_1^*, ..., s_{Nb}^*$, exceeding s_0 is an approximation of the p-value.

Gelper and Croux (2007) show that the test performs well in a Monte Carlo setting as well as in an application to real data.

3.A.5 Carter-Kohn Algorithm

This section lays out the Carter-Kohn algorithm. It is used to sample the factors Y given all model parameters, $\boldsymbol{\theta}$, and the data, X, i.e. it generates draws from $p(Y|\boldsymbol{\theta}, X)$.

State-space form

Start by rewriting observation and transition equation as

$$\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{z}_t \end{bmatrix} = \mathcal{H}\mathcal{B}_t + \mathcal{W}_t \tag{3.72}$$

$$\mathcal{B}_t = \mathcal{F}\mathcal{B}_{t-1} + \mathcal{V}_t \tag{3.73}$$

$$Var(\mathcal{W}_t) = \mathcal{R} \tag{3.74}$$

$$Var(\mathcal{V}_t) = \mathcal{Q} \tag{3.75}$$

where

$$\mathcal{H} = \begin{bmatrix} \Lambda^f & \Lambda^z & 0 & ... & 0 \\ \mathbf{0} & I & 0 & ... & 0 \end{bmatrix}; \quad \mathcal{B}_t = \begin{bmatrix} \beta_t' & \beta_{t-1}' & ... & \beta_{t-p}' \end{bmatrix}'; \quad \beta_t = \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{z}_t \end{bmatrix}$$

$$\mathcal{W}_t = \begin{bmatrix} \boldsymbol{\xi}_t \\ \mathbf{0} \end{bmatrix}; \quad \mathcal{F} = \begin{bmatrix} & \Pi & \\ I & \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathcal{V}_t = \begin{bmatrix} \boldsymbol{u}_t \\ \mathbf{0} \end{bmatrix}$$

$$\mathcal{R} = \begin{bmatrix} \Omega & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathcal{Q} = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Then consider the following factorisation:

$$p(\mathcal{B}_{1:T}|X,\boldsymbol{\theta}) = p(\mathcal{B}_T|\boldsymbol{x}_{1:T},\boldsymbol{\theta})\prod_{t=1}^{T-1} p(\mathcal{B}_t|\mathcal{B}_{t+1},X,\theta) \tag{3.76}$$

Given the linear Gaussian form of the state space model we have that

$$\mathcal{B}_T|\boldsymbol{x}_{1:T},\theta \sim N(\mathcal{B}_{T|T},\mathcal{P}_{T|T}) \tag{3.77}$$

$$\mathcal{B}_{t|T}|\mathcal{B}_{t+1|T},\boldsymbol{x}_{1:T},\theta \sim N(\mathcal{B}_{t|t,\mathcal{B}_{t+1|T}},\mathcal{P}_{t|t,\mathcal{B}_{t+1|T}}) \tag{3.78}$$

with

$$\mathcal{B}_{T|T} = E(\mathcal{B}_T|\boldsymbol{x}_{1:T},\boldsymbol{\theta}) \tag{3.79}$$

$$\mathcal{P}_{T|T} = Cov(\mathcal{B}_T|\boldsymbol{x}_{1:T},\boldsymbol{\theta}) \tag{3.80}$$

$$\mathcal{B}_{t|t,\mathcal{B}_{t+1|T}} = E(\mathcal{B}_t|\mathcal{B}_{t|t},\mathcal{B}_{t+1|t},\boldsymbol{\theta}) \tag{3.81}$$

$$\mathcal{P}_{t|t,\mathcal{B}_{t+1|T}} = Cov(\mathcal{B}_t|\mathcal{B}_{t|t},\mathcal{B}_{t+1|t},\boldsymbol{\theta}) \tag{3.82}$$

Kalman-filter

In a first step, I run a Kalman filter to obtain a series of Kalman-filtered draws of the state variable $\mathcal{B}_t$ $\mathcal{B}_{t|t}$ for $t = 1,...T$. To initialise, I set $\mathcal{B}_{1|0} = 0$ and $\mathcal{P}_{1|0} = \mathbf{I}$. Then, iterate forward as:

$$\mathcal{B}_{t|t} = \mathcal{B}_{t|t-1} + \kappa_{t|t-1}\eta_{t|t-1} \tag{3.83}$$

where $\eta_{t|t-1} = \mathcal{B}_t - \mathcal{F}\mathcal{B}_{t|t-1}$ denotes the forecast error, $\mathfrak{f}_{t|t-1} = \mathcal{H}\mathcal{P}_{t|t-1}\mathcal{H}' + \mathcal{R}$ its variance and $\kappa_{t|t-1} = \mathcal{P}_{t|t-1}\mathcal{H}\mathfrak{f}_{t|t-1}^{-1}$ the "Kalman-gain"

$$\mathcal{P}_{t|t-1} = \mathcal{F}\mathcal{P}_{t-1|t-1}\mathcal{F}' + \mathcal{Q} \tag{3.84}$$

Then, conditioning on the last of these Kalman-filtered draws, $\mathcal{B}_{T|T}$ and $\mathcal{P}_{T|T}$, run the filter backwards to obtain a series $\mathcal{B}_{t|t+1}$ for $t = 1,...,T-1$ as follows:

$$\mathcal{B}^*_{t|t,\mathcal{B}_{t+1|T}} = \mathcal{B}_{t|t} + \mathcal{P}_{t|t}\mathcal{F}^{*'}J^{-1}_{t+1|t}\psi_{t+1|t} \tag{3.85}$$

$$\mathcal{P}^*_{t|t,\mathcal{B}_{t+1|T}} = \mathcal{P}_{t|t} - \mathcal{P}_{t|t}\mathcal{F}^{*'}J^{-1}_{t+1|t}\mathcal{F}^*\mathcal{P}_{t|t} \tag{3.86}$$

where $\psi_{t+1|t} = \mathcal{B}^*_{t+1} - \mathcal{F}^*\mathcal{B}_{t|t}$ and $J_{t+1|t} = \mathcal{F}^*P_{t|t}\mathcal{F}^{*'} + \mathcal{Q}^*$. Note that $\mathcal{Q}^*$ refers to the top $R \times R$ block of $\mathcal{Q}$ and that $\mathcal{F}^*$ and $\mathcal{B}^*$ denote the first R rows of $\mathcal{F}$ and $\mathcal{B}$, respectively. This is required because $\mathcal{Q}$ is singular given the presence of observable factors.

Plugging these draws into (3.76) results in an unconditional posterior draw of the state variable, $\mathcal{B}_1, ..., \mathcal{B}_T$. Its top $R + K$ block represents an unconditional posterior draw of factors, $\boldsymbol{y}_t$.

Convergence of the Posterior Sampling Algorithm

The convergence properties of the reduced form parameters of a Bayesian FAVAR model are discussed in detail in Amir-Ahmadi and Uhlig (2015). They show that a Gibbs sampling procedure, similar to the one employed for the reduced form parameters here, converges for appropriate lengths of the MCMC chain. The convergence properties of the structural parameters, however, need to be assessed. In particular the first column of B containing the on-impact effects of the shock of interest are of importance. In order to do so, I follow Amir-Ahmadi and Uhlig (2015) and employ the convergence diagnostic proposed by Geweke (1992). A detailed discussion of this convergence diagnostic can be found, for example, in Cowles and Carlin (1996).

This diagnostic assesses the convergence of each element η_i of a parameter vector, $\boldsymbol{\eta}$. The assessment is based on a comparison of means across different parts of this chain. If the means are close to each other, the procedure detects convergence.

In a first step, extract from each (univariate) posterior draw $\{\eta_i\}_{i=1}^{D}$ the following subseries: $\eta_{1i}, ..., \eta_{0.1D,i}$, i.e. the first 10 % of draws for parameter i, and $\eta_{0.6D+1,i}, ..., \eta_{D,i}$, i.e. the last 40% of draws, where D is the length of the MCMC chain. Compute $\hat{\mu}_{first}$ and $\hat{\mu}_{last}$, the mean, as well as $\hat{\sigma}_{first}$ and $\hat{\sigma}_{last}$, the standard deviation, of these subseries. Then the test statistic is

$$CD = \frac{\hat{\mu}_{first} - \hat{\mu}_{last}}{\frac{\hat{\sigma}_{first}}{\sqrt{0.1D}} + \frac{\hat{\sigma}_{last}}{\sqrt{0.4D}}} \tag{3.87}$$

Under the conditions mentioned in Geweke (1992), CD has an asymptotic standard normal distribution

The final output is a p-value indicating whether or not we can reject the null hypothesis of convergence, i.e. equality of mean across the chain, at a given significance level.

Table 3.4 shows that for 10 out of 12 parameters, the null hypothesis of convergence cannot be rejected. Note that the first R parameters refer to the on-impact effect on the latent factors, which do not have an economic interpretation.

3.A.6 Criteria to determine the number of factors

This section describes two criteria to determine the number of factors: The Bai and Ng (2002) criterion and the scree plot.

Table 3.4: Geweke (1992) test for convergence of the MCMC chain

	p-value
r=1	0.7348
r=2	0.0007
r=3	0.1343
r=4	0.1363
k=1	0.8634
k=2	0.1635
k=3	0.7797
k=4	0.0871
k=5	0.3087
k=6	0.9196
k=7	0.0444
k=8	0.1515

Note: The table shows the test for convergence of the parameters reflecting the on-impact effects of the structural shocks. The first $r = 1, ..., 4$ values refer to the impact effect of the latent factors. The last $k = 1, ..., 8$ refer to the on-impact effects of observable variables. The sample split is $T = T_1 + T_2 + T_3$, where $T_1 = 0.1T$, $T_2 = 0.5T$ and $T_3 = 0.4T$. P-values are computed as quantiles of a Chi-squared distribution.

Bai and Ng (2002) criterion

Bai and Ng (2002) suggest the following criterion to determine the number of factors:

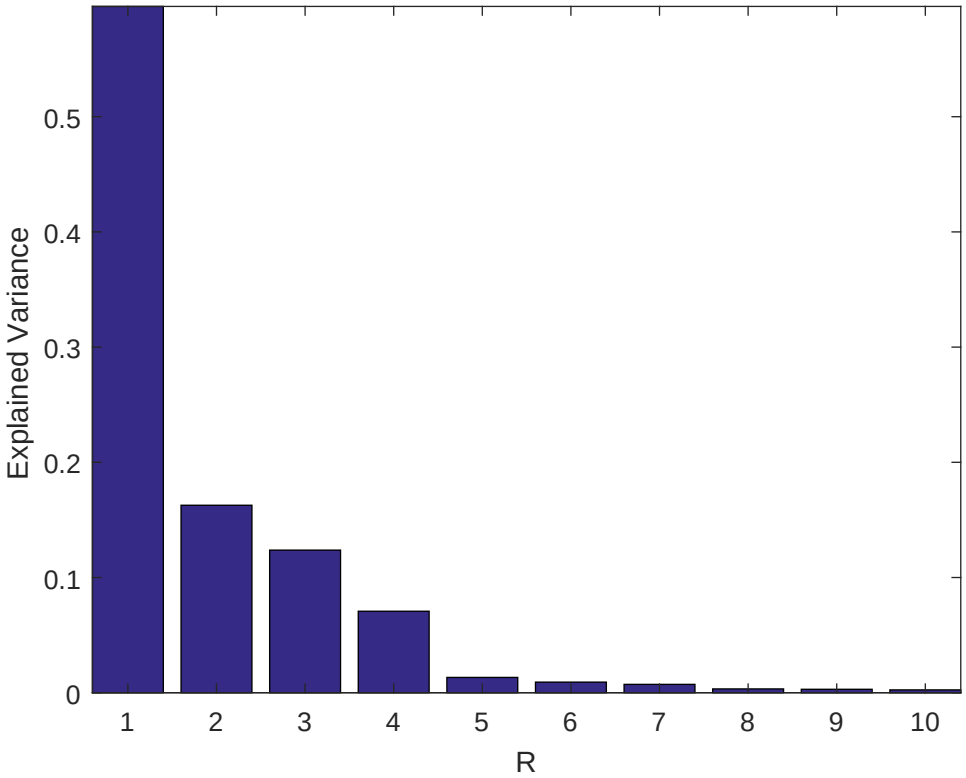
$$\begin{aligned} BN(R) =& \log \frac{1}{NT} \sum_{t=1}^{T} (X_t - \Lambda^{PC} \boldsymbol{f}_t^{PC})'(X_t - \Lambda^{PC} \boldsymbol{f}_t^{PC}) \\ &+ R \frac{N+T}{NT} log(min(N,T)) \end{aligned} \tag{3.88}$$

where Λ^{PC} and $\boldsymbol{f}_t^{PC}$ are the principal components estimators of the factor loadings and the factors, respectively. Bai and Ng (2002) suggest to set R^* such that (3.88) is minimised.

Scree Plot

A scree plot summarizes the marginal contribution of the r-th factor to the average explanatory power of N regressions of $\boldsymbol{x}_t$ against the first r factors as computed via Principal Components. Figure 3.14 plots this marginal contribution against the number of factors. It shows that the first factor explains about 60% of variance in $\boldsymbol{x}_t$, while the first four factors explain over 95% of the variance, suggesting that for $R = 4$ the factors explain a sufficiently large share of the variance in $\boldsymbol{x}_t$.

Figure 3.14: Scree Plot



Note: Explained share of variance in $\boldsymbol{x}_t$ as a function of the number of latent factors (R) included in $\boldsymbol{f}_t$

3.A.7 Data Description

Table 3.5: Data

Output and Income

id	tcode	fred	description
1	5	RPI	Real Personal Income
2	5	W875RX1	Real personal income ex transfer receipts
6	5	INDPRO	IP Index
7	5	IPFPNSS	IP: Final Products and Nonindustrial Supplies
8	5	IPFINAL	IP: Final Products (Market Group)
9	5	IPCONGD	IP: Consumer Goods
10	5	IPDCONGD	IP: Durable Consumer Goods
11	5	IPNCONGD	IP: Nondurable Consumer Goods
12	5	IPBUSEQ	IP: Business Equipment
13	5	IPMAT	IP: Materials
14	5	IPDMAT	IP: Durable Materials
15	5	IPNMAT	IP: Nondurable Materials
16	5	IPMANSICS	IP: Manufacturing (SIC)
17	5	IPB51222s	IP: Residential Utilities
18	5	IPFUELS	IP: Fuels
19	1	NAPMPI	ISM Manufacturing: Production Index
20	2	CUMFNS	Capacity Utilization: Manufacturing

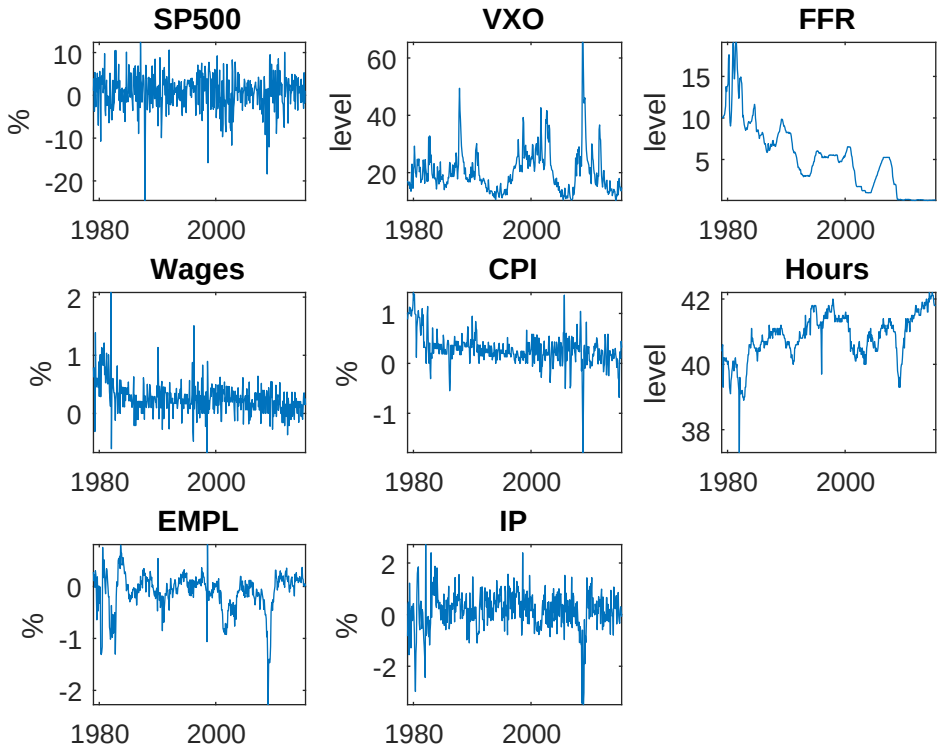
Labor Market			
id	tcode	fred	description
21*	2	HWI	Help-Wanted Index for United States
22*	2	HWIURATIO	Ratio of Help Wanted/No. Unemployed
23	5	CLF16OV	Civilian Labor Force
24	5	CE16OV	Civilian Employment
25	2	UNRATE	Civilian Unemployment Rate
26	2	UEMPMEAN	Average Duration of Unemployment (Weeks)
27	5	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks
28	5	UEMP5TO14	Civilians Unemployed for 5-14 Weeks
29	5	UEMP15OV	Civilians Unemployed - 15 Weeks & Over
30	5	UEMP15T26	Civilians Unemployed for 15-26 Weeks
31	5	UEMP27OV	Civilians Unemployed for 27 Weeks and Over
32*	5	CLAIMSx	Initial Claims
33	5	PAYEMS	All Employees: Total nonfarm
34	5	USGOOD	All Employees: Goods-Producing Industries
35	5	CES1021000001	All Employees: Mining and Logging: Mining
36	5	USCONS	All Employees: Construction
37	5	MANEMP	All Employees: Manufacturing
38	5	DMANEMP	All Employees: Durable goods
39	5	NDMANEMP	All Employees: Nondurable goods
40	5	SRVPRD	All Employees: Service-Providing Industries
41	5	USTPU	All Employees: Trade, Transportation & Utilities
42	5	USWTRADE	All Employees: Wholesale Trade
43	5	USTRADE	All Employees: Retail Trade
44	5	USFIRE	All Employees: Financial Activities
45	5	USGOVT	All Employees: Government
46	1	CES0600000007	Avg Weekly Hours : Goods-Producing
47	2	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing
48	1	AWHMAN	Avg Weekly Hours : Manufacturing
49	1	NAPMEI	ISM Manufacturing: Employment Index
127	6	CES0600000008	Avg Hourly Earnings : Goods-Producing
128	6	CES2000000008	Avg Hourly Earnings : Construction
129	6	CES3000000008	Avg Hourly Earnings : Manufacturing

Housing			
id	tcode	fred	description
50	4	HOUST	Housing Starts: Total New Privately Owned
51	4	HOUSTNE	Housing Starts, Northeast
52	4	HOUSTMW	Housing Starts, Midwest
53	4	HOUSTS	Housing Starts, South
54	4	HOUSTW	Housing Starts, West
55	4	PERMIT	New Private Housing Permits (SAAR)
56	4	PERMITNE	New Private Housing Permits, Northeast (SAAR)
57	4	PERMITMW	New Private Housing Permits, Midwest (SAAR)
58	4	PERMITS	New Private Housing Permits, South (SAAR)
59	4	PERMITW	New Private Housing Permits, West (SAAR)
Consumption, Orders and inventories			
3	5	DPCERA3M086SBEA	Real personal consumption expenditures
4*	5	CMRMTSPLx	Real Manu. and Trade Industries Sales
5*	5	RETAILx	Retail and Food Services Sales
60	1	NAPM	ISM : PMI Composite Index
61	1	NAPMNOI	ISM : New Orders Index
62	1	NAPMSDI	ISM : Supplier Deliveries Index
63	1	NAPMII	ISM : Inventories Index
64	5	ACOGNO	New Orders for Consumer Goods
65*	5	AMDMNOx	New Orders for Durable Goods
66*	5	ANDENOx	New Orders for Nondefense Capital Goods
67*	5	AMDMUOx	Unfilled Orders for Durable Goods
68*	5	BUSINVx	Total Business Inventories
69*	2	ISRATIOx	Total Business: Inventories to Sales Ratio
130*	2	UMCSENTx	Consumer Sentiment Index

id	tcode	fred	description
Money and Credit			
70	6	M1SL	M1 Money Stock
71	6	M2SL	M2 Money Stock
72	5	M2REAL	Real M2 Money Stock
73	6	AMBSL	St. Louis Adjusted Monetary Base
74	6	TOTRESNS	Total Reserves of Depository Institutions
75	7	NONBORRES	Reserves Of Depository Institutions
76	6	BUSLOANS	Commercial and Industrial Loans
77	6	REALLN	Real Estate Loans at All Commercial Banks
78	6	NONREVSL	Total Nonrevolving Credit
79*	2	CONSPI	Nonrevolving consumer credit to Personal Income
131	6	MZMSL	MZM Money Stock
132	6	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding
133	6	DTCTHFNM	Total Consumer Loans and Leases Outstanding
134	6	INVEST	Securities in Bank Credit at All Commercial Banks
Interest Rates and Exchange Rates			
84	2	FEDFUNDS	Effective Federal Funds Rate
85*	2	CP3Mx	3-Month AA Financial Commercial Paper Rate
86	2	TB3MS	3-Month Treasury Bill:
87	2	TB6MS	6-Month Treasury Bill:
88	2	GS1	1-Year Treasury Rate
89	2	GS5	5-Year Treasury Rate
90	2	GS10	10-Year Treasury Rate
91	2	AAA	Moody's Seasoned Aaa Corporate Bond Yield
92	2	BAA	Moody's Seasoned Baa Corporate Bond Yield
93*	1	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS
94	1	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
95	1	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
96	1	T1YFFM	1-Year Treasury C Minus FEDFUNDS
97	1	T5YFFM	5-Year Treasury C Minus FEDFUNDS
98	1	T10YFFM	10-Year Treasury C Minus FEDFUNDS
99	1	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
100	1	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
101	5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies
102*	5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
103*	5	EXJPUSx	Japan / U.S. Foreign Exchange Rate
104*	5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate
105*	5	EXCAUSx	Canada / U.S. Foreign Exchange Rate

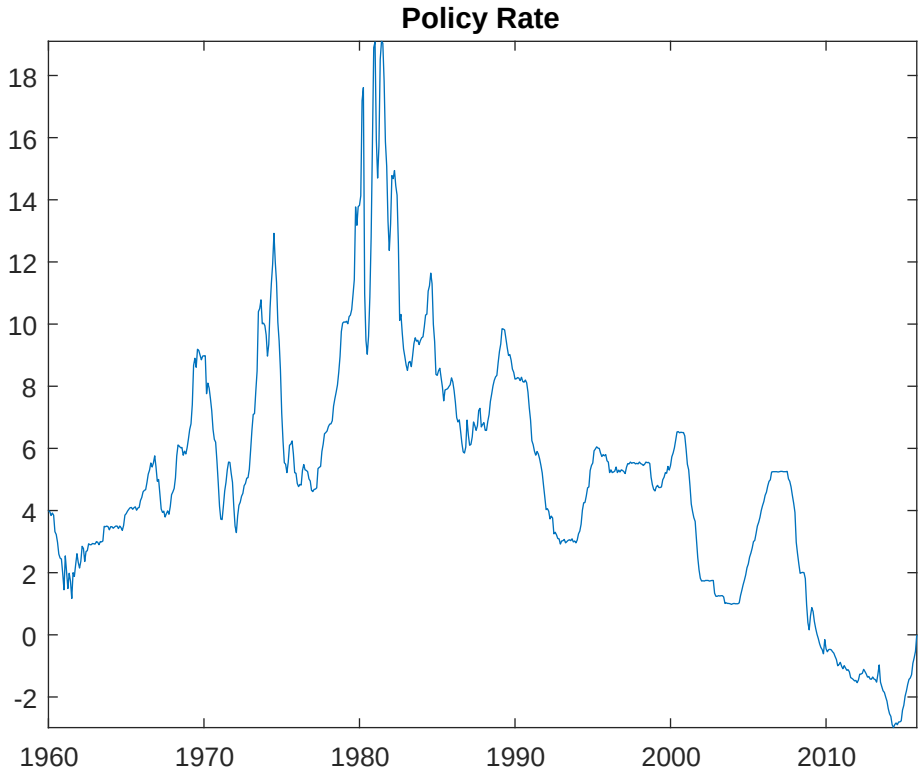
Prices			
id	tcode	fred	description
106	6	WPSFD49207	PPI: Finished Goods
107	6	WPSFD49502	PPI: Finished Consumer Goods
108	6	WPSID61	PPI: Intermediate Materials
109	6	WPSID62	PPI: Crude Materials
110*	6	OILPRICEx	Crude Oil, spliced WTI and Cushing
111	6	PPICMM	PPI: Metals and metal products:
112	1	NAPMPRI	ISM Manufacturing: Prices Index
113	6	CPIAUCSL	CPI : All Items
114	6	CPIAPPSL	CPI : Apparel
115	6	CPITRNSL	CPI : Transportation
116	6	CPIMEDSL	CPI : Medical Care
117	6	CUSR0000SAC	CPI : Commodities
118	6	CUUR0000SAD	CPI : Durables
119	6	CUSR0000SAS	CPI : Services
120	6	CPIULFSL	CPI : All Items Less Food
121	6	CUUR0000SA0L2	CPI : All items less shelter
122	6	CUSR0000SA0L5	CPI : All items less medical care
123	6	PCEPI	Personal Cons. Expend.: Chain Index
124	6	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods
125	6	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods
126	6	DSERRG3M086SBEA	Personal Cons. Exp: Services
Stock Market			
80*	5	S&P 500	S&P's Common Stock Price Index: Composite
81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials
82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield
83*	5	S&P PE ratio	S&P's Composite Common Stock: PE Ratio
135*	1	VXOCLSx	VXO

Figure 3.15: Observable Factors



Note: The sample length is 1979M1 to 2015M7. The observable factors are the variables included in Piffer and Podstawski (2017).

Figure 3.16: Shadow Rate



Note: Wu and Xia (2016) shadow rate available from 1960M1 to 2015M11.

Bibliography

Alberola, E., Estrada, Á. and Santabárbara, D. (2013), 'Growth beyond imbalances. Sustainable growth rates and output gap reassessment', *No. 1313. Banco de España Documentos de Trabajo* .

Alesina, A., Prati, A. and Tabellini, G. (1989), 'Public confidence and debt management: A model and a case study of Italy', *No. 351. CEPR Discussion Papers* .

Alessi, L., Antunes, A., Babeckỳ, J., Baltussen, S., Behn, M., Bonfim, D., Bush, O., Detken, C., Frost, J., Guimaraes, R. et al. (2015), 'Comparing different early warning systems: Results from a horse race competition among members of the macro-prudential research network', *No. 62194. University Library of Munich* .

Amir-Ahmadi, P. and Drautzburg, T. (2018), 'Identification and inference with ranking restrictions', *FRB Philadelphia working paper No.17-11* .

Amir-Ahmadi, P. and Uhlig, H. (2015), 'Sign restrictions in Bayesian FAVARs with an application to monetary policy shocks', *NBER Working paper 21738* .

Angelini, G., Bacchiocchi, E., Caggiano, G. and Fanelli, L. (2019), 'Uncertainty across volatility regimes', *Journal of Applied Econometrics* **34**(3), 437–455.

Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018), 'Narrative sign restrictions for SVARs', *American Economic Review* **108**(10), 2802–29.

Arias, J. E., Caldara, D. and Rubio-Ramirez, J. F. (2019), 'The systematic component of monetary policy in SVARs: an agnostic identification procedure', *Journal of Monetary Economics* **101**, 1–13.

Arias, J. E., Rubio-Ramírez, J. F. and Waggoner, D. F. (2018), 'Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications', *Econometrica* **86**(2), 685–720.

Bachmann, R. and Bayer, C. (2013), "Wait-and-see' business cycles?', *Journal of Monetary Economics* **60**(6), 704–719.

Bachmann, R. and Sims, E. R. (2012), 'Confidence and the transmission of government spending shocks', *Journal of Monetary Economics* **59**(3), 235–249.

Bai, J. and Ng, S. (2002), 'Determining the number of factors in approximate factor models', *Econometrica* **70**(1), 191–221.

Baker, S. R., Bloom, N. and Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Baldacci, E., Petrova, I. K., Belhocine, N., Dobrescu, G. and Mazraani, S. (2011), 'Assessing fiscal stress', *IMF Working Papers No. 11/100.* pp. 1–41.

Barnichon, R. and Matthes, C. (2018), 'Functional approximation of impulse responses', *Journal of Monetary Economics* **99**, 41–55.

Basu, S. and Bundick, B. (2017), 'Uncertainty shocks in a model of effective demand', *Econometrica* **85**(3), 937–958.

Baumeister, C. and Hamilton, J. D. (2015), 'Sign restrictions, structural vector autoregressions, and useful prior information', *Econometrica* **83**(5), 1963–1999.

Baumeister, C. J. and Hamilton, J. D. (2018), 'Inference in structural vector autoregressions when the identifying assumptions are not fully believed: Re-evaluating the role of monetary policy in economic fluctuations', *Journal of Monetary Economics* **100**, 48–65.

Baumeister, C. J. and Hamilton, J. D. (2019), 'Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks', *The American Economic Review* **109**(5), 1873–1910.

Belviso, F. and Milani, F. (2006), 'Structural factor-augmented VARs (SFAVARs) and the effects of monetary policy', *Topics in Macroeconomics* **6**(3).

Benati, L. and Surico, P. (2009), 'VAR analysis and the great moderation', *American Economic Review* **99**(4), 1636–52.

Berg, A., Borensztein, E. and Pattillo, C. (2005), 'Assessing early warning systems: how have they worked in practice?', *IMF staff papers* **52**(3), 462–502.

Bernanke, B. S., Boivin, J. and Eliasz, P. (2005), 'Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach', *The Quarterly Journal of Economics* **120**(1), 387–422.

Berti, K., Salto, M., Lequien, M. et al. (2012), 'An early-detection index of fiscal stress for EU countries', *Directorate General Economic and Financial Affairs (DG ECFIN), European Commission* .

Bibby, J., Kent, J. and Mardia, K. (1979), *Multivariate analysis*, Academic Press, London.

Binning, A. (2013), 'Underidentified SVAR models: A framework for combining short and long-run restrictions with sign restrictions', *Norges Bank working paper No. 2013/14* .

Bloom, N. (2009), 'The impact of uncertainty shocks', *Econometrica* **77**(3), 623–685.

Boivin, J., Giannoni, M. P. and Mihov, I. (2009), 'Sticky prices and monetary policy: Evidence from disaggregated US data', *The American Economic Review* **99**(1), 350–384.

Borio, C., Disyatat, P. and Juselius, M. (2016), 'Rethinking potential output: Embedding information about the financial cycle', *Oxford Economic Papers* **69**(3), 655–677.

Born, B. and Pfeifer, J. (2014), 'Policy risk and the business cycle', *Journal of Monetary Economics* **68**, 68–85.

Braun, R., Brüggemann, R. et al. (2017), 'Identification of SVAR models by combining sign restrictions with external instruments', *Working Paper No. 2017-07, Department of Economics, University of Konstanz* .

Brooks, S. P. and Gelman, A. (1998), 'General methods for monitoring convergence of iterative simulations', *Journal of computational and graphical statistics* **7**(4), 434–455.

Brooks, S. P. and Roberts, G. O. (1998), 'Assessing convergence of Markov Chain Monte Carlo algorithms', *Statistics and Computing* **8**(4), 319–335.

Bussiere, M. (2013), 'Balance of payment crises in emerging markets: how early were the early warning signals?', *Applied Economics* **45**(12), 1601–1623.

Caggiano, G., Castelnuovo, E. and Nodari, G. (2017), 'Uncertainty and monetary policy in good and bad times', *No. 6630. CESifo Working Paper* .

Caggiano, G., Castelnuovo, E. and Pellegrino, G. (2017), 'Estimating the real effects of uncertainty shocks at the zero lower bound', *European Economic Review* **100**, 257–272.

Caldara, D., Cavallo, M. and Iacoviello, M. (2018), 'Oil price elasticities and oil price fluctuations', *Journal of Monetary Economics* **103**, 1–20.

Caldara, D. and Herbst, E. (2019), 'Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs', *American Economic Journal: Macroeconomics* **11**(1), 157–92.

Calvo, G. A. (1988), 'Servicing the public debt: The role of expectations', *The American Economic Review* **78**(4), 647–661.

Canova, F. (2007), *Methods for applied macroeconomic research*, Vol. 13, Princeton University Press.

Canova, F. and De Nicoló, G. (2002), 'Monetary disturbances matter for business fluctuations in the G-7', *Journal of Monetary Economics* **49**(6), 1131–1159.

Canova, F. and Paustian, M. (2011), 'Business cycle measurement with some theory', *Journal of Monetary Economics* **58**(4), 345–361.

Canova, F. and Pina, J. P. (2005), What VAR tell us about DSGE models?, *in* 'New Trends in Macroeconomics', Springer, pp. 89–123.

Carriero, A., Clark, T. E. and Marcellino, M. (2018), 'Measuring uncertainty and its impact on the economy', *Review of Economics and Statistics* **100**(5), 799–815.

Carter, C. K. and Kohn, R. (1994), 'On Gibbs sampling for state space models', *Biometrika* **81**(3), 541–553.

Chakrabarti, A. and Zeaiter, H. (2014), 'The determinants of sovereign default: A sensitivity analysis', *International Review of Economics & Finance* **33**, 300–318.

Christiano, L. J. and Fitzgerald, T. J. (2003), 'The band pass filter', *International Economic Review* **44**(2), 435–465.

Christiano, L. J., Motto, R. and Rostagno, M. (2014), 'Risk shocks', *American Economic Review* **104**(1), 27–65.

Christofides, C., Eicher, T. S. and Papageorgiou, C. (2016), 'Did established early warning signals predict the 2008 crises?', *European Economic Review* **81**, 103–114.

Cole, H. L. and Kehoe, T. J. (1996), 'A self-fulfilling model of mexico's 1994–1995 debt crisis', *Journal of International Economics* **41**(3-4), 309–330.

Comelli, F. (2013), 'Comparing parametric and non-parametric early warning systems for currency crises in emerging market economies', *IMF Working Paper No. 13/134.* .

Cowles, M. K. and Carlin, B. P. (1996), 'Markov Chain Monte Carlo convergence diagnostics: a comparative review', *Journal of the American Statistical Association* **91**(434), 883–904.

Creal, D. (2012), 'A survey of sequential Monte Carlo methods for economics and finance', *Econometric reviews* **31**(3), 245–296.

d Abiad, A. et al. (2003), 'Early warning systems; a survey and a regime-switching approach', *IMF Working Paper No. 03/32* .

De Cos, P. H., Moral-Benito, E., Koester, G. B. and Nickel, C. (2014), 'Signalling fiscal stress in the euro area: A country-specific early warning system', *Banco de Espana Working Paper No. 1418* .

Detragiache, M. E. and Spilimbergo, M. A. (2001), *Crises and liquidity: evidence and interpretation*, number 1-2, International Monetary Fund.

Eichengreen, B., Rose, A. K. and Wyplosz, C. (1995), 'Exchange market mayhem: the antecedents and aftermath of speculative attacks', *Economic policy* **10**(21), 249–312.

Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K. and Rubio-Ramírez, J. (2015), 'Fiscal volatility shocks and economic activity', *The American Economic Review* **105**(11), 3352–3384.

Forni, M. and Gambetti, L. (2014), 'Sufficient information in structural VARs', *Journal of Monetary Economics* **66**, 124–136.

Frankel, J. and Saravelos, G. (2012), 'Can leading indicators assess country vulnerability? Evidence from the 2008–09 global financial crisis', *Journal of International Economics* **87**(2), 216–231.

Frühwirth-Schnatter, S. (1994), 'Data augmentation and dynamic linear models', *Journal of time series analysis* **15**(2), 183–202.

Fry, R. and Pagan, A. (2011), 'Sign restrictions in structural vector autoregressions: A critical review', *Journal of Economic Literature* **49**(4), 938–960.

Gelman, A. and Rubin, D. B. (1992), 'Inference from iterative simulation using multiple sequences', *Statistical science* pp. 457–472.

Gelper, S. and Croux, C. (2007), 'Multivariate out-of-sample tests for granger causality', *Computational statistics & data analysis* **51**(7), 3319–3329.

Geman, S. and Geman, D. (1984), 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images', *IEEE Transactions on pattern analysis and machine intelligence* (6), 721–741.

Geweke, J. (1989), 'Bayesian inference in econometric models using Monte Carlo integration', *Econometrica* **57**(6), 1317–1339.

Geweke, J. (1992), 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', *Bayesian Statistics* **4**, 169–193.

Giacomini, R. and Kitagawa, T. (2015), 'Robust inference about partially identified SVARs', *Manuscript, UCL* .

Gourinchas, P.-O. and Obstfeld, M. (2012), 'Stories of the twentieth century for the twenty-first', *American Economic Journal: Macroeconomics* **4**(1), 226–65.

Herbst, E. and Schorfheide, F. (2014), 'Sequential Monte Carlo sampling for DSGE models', *Journal of Applied Econometrics* **29**(7), 1073–1098.

Ho, T.-k. (2015), 'Looking for a needle in a haystack: Revisiting the cross-country causes of the 2008–9 crisis by Bayesian model averaging', *Economica* **82**(328), 813–840.

Huber, F. and Fischer, M. M. (2018), 'A Markov switching factor-augmented VAR model for analyzing US business cycles and monetary policy', *Oxford Bulletin of Economics and Statistics* **80**(3), 575–604.

IMF (2010), 'The imf-fsb early warning exercise–design and methodological toolkit', *IMF Policy Paper, September* .

Jeanne, O. (1997), 'Are currency crises self-fulfilling?: A test', *Journal of International Economics* **43**(3-4), 263–286.

Jurado, K., Ludvigson, S. C. and Ng, S. (2015), 'Measuring uncertainty', *The American Economic Review* **105**(3), 1177–1216.

Kadiyala, K. R. and Karlsson, S. (1997), 'Numerical methods for estimation and inference in Bayesian VAR-models', *Journal of Applied Econometrics* pp. 99–132.

Kaminsky, G. L. and Reinhart, C. M. (1999), 'The twin crises: the causes of banking and balance-of-payments problems', *American Economic Review* **89**(3), 473–500.

Kaminsky, G., Lizondo, S. and Reinhart, C. M. (1998), 'Leading indicators of currency crises', *Staff Papers-International Monetary Fund* pp. 1–48.

Kilian, L. (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *American Economic Review* **99**(3), 1053–69.

Kilian, L. and Lütkepohl, H. (2017), *Structural vector autoregressive analysis*, Cambridge University Press.

Kilian, L. and Murphy, D. P. (2012), 'Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models', *Journal of the European Economic Association* **10**(5), 1166–1188.

Klau, M., Hawkins, J. et al. (2000), 'Measuring potential vulnerabilities in emerging market economies', *Bank for International Settlements Working Paper, No.91* .

Kociecki, A. (2010), 'A prior for impulse responses in Bayesian structural VAR models', *Journal of Business & Economic Statistics* **28**(1), 115–127.

Koop, G. (2003), *Bayesian Econometrics*, John Wiley & Sons Ltd,.

Koop, G., Korobilis, D. et al. (2010), 'Bayesian multivariate time series methods for empirical macroeconomics', *Foundations and Trends in Econometrics* **3**(4), 267–358.

Koopman, S. J., Shephard, N. and Creal, D. (2009), 'Testing the assumptions behind importance sampling', *Journal of Econometrics* **149**(1), 2–11.

Kraay, A. and Nehru, V. (2006), 'When is external debt sustainable?', *The World Bank Economic Review* **20**(3), 341–365.

Krugman, P. (1979), 'A model of balance-of-payments crises', *Journal of money, credit and banking* **11**(3), 311–325.

Leamer, E. E. (1985), 'Sensitivity analyses would help', *The American Economic Review* **75**(3), 308–313.

Leduc, S. and Liu, Z. (2016), 'Uncertainty shocks are aggregate demand shocks', *Journal of Monetary Economics* **82**, 20–35.

Litterman, R. B. (1986), 'Forecasting with Bayesian vector autoregressions: five years of experience', *Journal of Business & Economic Statistics* **4**(1), 25–38.

Ludvigson, S. C., Ma, S. and Ng, S. (2018), 'Uncertainty and business cycles: exogenous impulse or endogenous response?', *NBER Working paper 21803* .

Lütkepohl, H. (2005), *New introduction to multiple time series analysis*, Springer Science & Business Media.

Manasse, P., Roubini, N. and Schimmelpfennig, A. (2003), 'Predicting sovereign debt crises', *IMF Working Paper No. 03/221* .

Mandalinci, Z. and Mumtaz, H. (2019), 'Global economic divergence and portfolio capital flows to emerging markets', *Journal of Money, Credit and Banking* .

Marashaden, O. (1997), 'A logit model to predict debt rescheduling by less developed countries', *Asian Economies* **26**, 25–34.

Masson, P. (1999), 'Multiple equilibria, contagion, and the emerging market crises', *IMF Working Paper No. 99/164* .

Mathai, A. M. and Haubold, H. J. (2008), *Special functions for applied scientists*, Vol. 4, Springer.

McCracken, M. W. and Ng, S. (2016), 'Fred-md: A monthly database for macroeconomic research', *Journal of Business & Economic Statistics* **34**(4), 574–589.

Obstfeld, M. et al. (1984), 'Balance-of-payments crises and devaluation', *Journal of Money, Credit and Banking* **16**(2), 208–217.

Peter, M. (2002), 'Estimating default probabilities of emerging market sovereigns: A new look at a not-so-new literature', *Working Paper Economics Section, The Graduate Institute of International Studies No. 06-2002* .

Pickands, J. et al. (1975), 'Statistical inference using extreme order statistics', *the Annals of Statistics* **3**(1), 119–131.

Piffer, M. and Podstawski, M. (2017), 'Identifying uncertainty shocks using the price of gold', *The Economic Journal* **128**(616), 3266–3284.

Plagborg-Møller, M. (2019), 'Bayesian inference on structural impulse response functions', *Quantitative Economics* **10**(1), 145–184.

Primiceri, G. E. (2005), 'Time varying structural vector autoregressions and monetary policy', *The Review of Economic Studies* **72**(3), 821–852.

Raftery, A. and Lewis, S. (1992), 'How many iterations in the Gibbs sampler?', *Bayesian Statistics* **4**, 763–773.

Reinhart, C. M. (2002), 'Default, currency crises and sovereign credit ratings', *NBER Working paper 8738* .

Reinhart, C. M. and Rogoff, K. S. (2008), 'This time is different: A panoramic view of eight centuries of financial crises', *NBER Working paper 13882* .

Robert, C. and Casella, G. (2013), *Monte Carlo statistical methods*, Springer Science & Business Media.

Rubio-Ramirez, J. F., Waggoner, D. F. and Zha, T. (2010), 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies* **77**(2), 665–696.

Sala-i Martin, X. (1997), 'I just ran two million regressions', *American Economic Review* **87**(2), 178–183.

Sims, C. A. and Zha, T. (1998), 'Bayesian methods for dynamic multivariate models', *International Economic Review* pp. 949–968.

Sims, E. R. (2012), News, non-invertibility, and structural VARs, *in* 'DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments', Emerald Group Publishing Limited, pp. 81–135.

Stock, J. H. and Watson, M. W. (2012), 'Disentangling the channels of the 2007-2009 recession', *NBER Working paper 18094* .

Stock, J. H. and Watson, M. W. (2016), 'Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics', *Handbook of macroeconomics* **2**, 415–525.

Sumner, S. P., Berti, K. et al. (2017), 'A complementary tool to monitor fiscal stress in european economies', *European Commission Discussion Paper 049* .

Uhlig, H. (2005), 'What are the effects of monetary policy on output? Results from an agnostic identification procedure', *Journal of Monetary Economics* **52**(2), 381–419.

Waggoner, D. F., Wu, H. and Zha, T. (2016), 'Striated Metropolis–Hastings sampler for high-dimensional models', *Journal of Econometrics* **192**(2), 406–420.

Wu, J. C. and Xia, F. D. (2016), 'Measuring the macroeconomic impact of monetary policy at the zero lower bound', *Journal of Money, Credit and Banking* **48**(2-3), 253–291.

Yamamoto, Y. (2019), 'Bootstrap inference for impulse response functions in factor-augmented vector autoregressions', *Journal of Applied Econometrics* **34**(2), 247–267.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorgelegte Dissertation auf Grundlage der angegebenen Quellen und Hilfsmittel selbstständig verfasst habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sind als solche kenntlich gemacht. Die vorgelegte Dissertation hat weder in der gleichen noch einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule zum Promotionsverfahren vorgelegen.

Martin Bruns
Berlin, den 13. Mai 2019

Liste verwendeter Hilfsmittel

- Matlab R2016b
  - Optimization Toolbox
  - Statistics Toolbox
  - Parallel Computing Toolbox
- Microsoft Excel
- Stata 13
- Fortran
- LaTeX
- Siehe auch Literatur- und Quellenangaben